

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2003, 04:02:09 ; Search time 1517 Seconds  
(without alignments)  
7712.692 Million cell updates/sec

Title: US-09-928-457-95  
Perfect score: 286  
Sequence: 1 AATCGGAGGACGATACCG.....GCTCCACCACTTCGGGAAAT 286

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.:

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_cm.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_man.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	286	100.0	286	1	AF169473	AF169473 Neisseria	
2	286	100.0	286	6	A68924	A68924 Sequence 95	
3	286	100.0	286	6	BD063024	BD063024 DNA and s	
4	177	61.9	3204	1	NME311654	AX311654 Neisseria	
5	177	61.9	3204	6	AX024064	AX024064 Sequence	
c	177	61.9	349061	1	NMA222491	AX162753 Neisseria	
7	64	22.4	3252	1	AY150284	AY150284 Neisseria	
8	64	22.4	3254	1	NME277537	AX277537 Neisseria	
9	64	22.4	3939	6	AX236409	AX236409 Sequence	
10	64	22.4	3939	6	AX236419	AX236419 Sequence	
11	64	22.4	3939	6	AX239717	AX239717 Sequence	
12	64	22.4	3939	6	AX239729	AX239729 Sequence	
13	64	22.4	4170	6	AX236463	AX236463 Sequence	
14	64	22.4	4170	6	AX239749	AX239749 Sequence	
15	64	22.4	4179	6	AX236413	AX236413 Sequence	
16	64	22.4	4179	6	AX239721	AX239721 Sequence	
17	64	22.4	4218	6	AX236469	AX236469 Sequence	
18	64	22.4	4218	6	AX239755	AX239755 Sequence	
19	64	22.4	4335	6	AX236457	AX236457 Sequence	
20	64	22.4	4335	6	AX239743	AX239743 Sequence	
21	64	22.4	4344	6	AX236411	AX236411 Sequence	
22	64	22.4	4344	6	AX239719	AX239719 Sequence	
23	64	22.4	4425	6	AX236407	AX236407 Sequence	
24	64	22.4	4425	6	AX239715	AX239715 Sequence	
25	64	22.4	10591	1	AE002545	AE002545 Neisseria	
26	64	22.4	349980	6	AX044034	AX044034 Sequence	
c	27	21	7.3	600	1	AF169448	AF169448 Neisseria
c	28	20	7.0	177182	10	AL669946	AL669946 Mouse DNA
29	20	7.0	188476	2	AC120504	AC120504 Didelphis	
30	20	7.0	191671	10	AL672249	AL672249 Mouse DNA	
31	20	7.0	216608	5	AL844559	AL844559 Zebrafish	
32	20	7.0	223455	2	AC109844	AC109844 Rattus no	
33	20	7.0	238533	2	AC132968	AC132968 Rattus no	
c	34	20	7.0	248679	2	BX530409	BX530409 Danio rer
35	19	6.6	699	11	BV042099	BV042099 S212P603	
c	36	19	6.6	3459	1	AF269511	AF269511 Staphyloc
c	37	19	6.6	3459	6	AX144831	AX144831 Sequence
38	19	6.6	12155	1	AE015876	AE015876 Shewanell	
c	39	19	6.6	37918	9	AP005232	AP005232 Homo sapi
c	40	19	6.6	40327	9	AP005379	AP005379 Homo sapi
41	19	6.6	61144	2	AC009931	AC009931 Homo sapi	
42	19	6.6	106763	9	AP002091	AP002091 Homo sapi	
43	19	6.6	126169	9	AC010291	AC010291 Homo sapi	
c	44	19	6.6	127946	8	AC104473	AC104473 Oryza sat
45	19	6.6	150892	5	AL929190	AL929190 Zebrafish	

# ALIGNMENTS

RESULT 1  
LOCUS AF169473 286 bp DNA linear BCT 09-AUG-2000  
DEFINITION Neisseria meningitidis strain Z2491 clone Em085 unknown sequence.  
ACCESSION AF169473  
VERSION AF169473.1 GI:9754681  
KEYWORDS  
SOURCE  
ORGANISM Neisseria meningitidis  
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
Neisseriaceae; Neisseria.  
1 (bases 1 to 286)  
REFERENCE  
AUTHORS Perrin, A., Nassif, X. and Tinsley, C. R.  
TITLE Identification of regions of the chromosome of Neisseria meningitidis and Neisseria gonorrhoeae which are specific to

pathogenic Neisseriae  
 Unpublished  
 2 (bases 1 to 286)  
 Perrin,A., Nassif,X. and Tinsley,C.R.  
 Direct Submission  
 Submitted (16-JUL-1999) Necker Medicine Faculty, INSERM U411, 156  
 rue de Vaugirard, Paris 75015, France  
 Location/Qualifiers  
 1. .286  
 /organism="Neisseria meningitidis"  
 /mol\_type="genomic DNA"  
 /strain="Z2491"  
 /db\_xref="taxon:487"  
 /clone="Em085"  
 BASE COUNT 88 a 73 c 63 g 62 t  
 ORIGIN  
 Query Match 100.0%; Score 286; DB 1; Length 286;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-147;  
 Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AATTCGAGGAGCAGTACCGCAAGCGTTGCTCGCTATTCGCGGGTGATATAACAGAC 60  
 Db |||||  
 QY 1 AATTCGAGGAGCAGTACCGCAAGCGTTGCTCGCTATTCGCGGGTGATATAACAGAC 60  
 Db |||||  
 QY 61 GAGGGTATCCGCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 120  
 Db |||||  
 QY 61 GAGGGTATCCGCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 120  
 Db |||||  
 QY 121 AAAAAATGCTTTTTCATTTTTCGCAAGCAATGACGACAGCTCAGCCCAACCAACT 180  
 Db |||||  
 QY 121 AAAAAATGCTTTTTCATTTTTCGCAAGCAATGACGACAGCTCAGCCCAACCAACT 180  
 Db |||||  
 QY 181 GACCCATTGCGCATTTTATGAAAAGACGCTCAAAAAGGCATTATCAGATTGCGGCGT 240  
 Db |||||  
 QY 181 GACCCATTGCGCATTTTATGAAAAGACGCTCAAAAAGGCATTATCAGATTGCGGCGT 240  
 Db |||||  
 QY 241 AGACCCAGTGGAGAAAAGTTCAATGGCTCCAAACCATTCGCGAATT 286  
 Db |||||  
 QY 241 AGACCCAGTGGAGAAAAGTTCAATGGCTCCAAACCATTCGCGAATT 286  
 Db |||||  
 RESULT 2  
 A68924  
 LOCUS  
 DEFINITION  
 Sequence 95 from Patent WO9802547.  
 ACCESSION  
 A68924  
 VERSION  
 A68924.1 GI:4759843  
 KEYWORDS  
 SOURCE  
 unidentified  
 ORGANISM  
 unclassified.  
 REFERENCE  
 1 (bases 1 to 286)  
 Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and  
 Merker,P.  
 TITLE  
 DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS  
 SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL  
 APPLICATIONS  
 JOURNAL  
 Patent: WO 9802547-A 95 22-JAN-1998;  
 INST NAT SANTE RECH MED (FR)  
 COMMENT  
 Other publication FR 2751000 19980116.  
 Location/Qualifiers  
 1. .286  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"  
 BASE COUNT 88 a 73 c 63 g 62 t  
 ORIGIN  
 Query Match 100.0%; Score 286; DB 6; Length 286;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-147;  
 Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AATTCGAGGAGCAGTACCGCAAGCGTTGCTCGCTATTCGCGGGTGATATAACAGAC 60  
 Db |||||  
 QY 1 AATTCGAGGAGCAGTACCGCAAGCGTTGCTCGCTATTCGCGGGTGATATAACAGAC 60  
 Db |||||  
 QY 61 GAGGGTATCCGCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 120  
 Db |||||  
 QY 61 GAGGGTATCCGCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 120  
 Db |||||  
 QY 121 AAAAAATGCTTTTTCATTTTTCGCAAGCAATGACGACAGCTCAGCCCAACCAACT 180  
 Db |||||  
 QY 121 AAAAAATGCTTTTTCATTTTTCGCAAGCAATGACGACAGCTCAGCCCAACCAACT 180  
 Db |||||  
 QY 181 GACCCATTGCGCATTTTATGAAAAGACGCTCAAAAAGGCATTATCAGATTGCGGCGT 240  
 Db |||||  
 QY 181 GACCCATTGCGCATTTTATGAAAAGACGCTCAAAAAGGCATTATCAGATTGCGGCGT 240  
 Db |||||  
 QY 241 AGACCCAGTGGAGAAAAGTTCAATGGCTCCAAACCATTCGCGAATT 286  
 Db |||||  
 QY 241 AGACCCAGTGGAGAAAAGTTCAATGGCTCCAAACCATTCGCGAATT 286  
 Db |||||  
 RESULT 2  
 A68924  
 LOCUS  
 DEFINITION  
 Sequence 95 from Patent WO9802547.  
 ACCESSION  
 A68924  
 VERSION  
 A68924.1 GI:4759843  
 KEYWORDS  
 SOURCE  
 unidentified  
 ORGANISM  
 unclassified.  
 REFERENCE  
 1 (bases 1 to 286)  
 Nassif,X., Tinsley,C., Achtman,M., Ruelle,J., Vinals,C. and  
 Merker,P.  
 TITLE  
 DNA AND SPECIFIC PROTEINS OR PEPTIDES OF THE NEISSERIA MENINGITIDIS  
 SPECIES BACTERIA, METHOD FOR OBTAINING THEM AND THEIR BIOLOGICAL  
 APPLICATIONS  
 JOURNAL  
 Patent: WO 9802547-A 95 22-JAN-1998;  
 INST NAT SANTE RECH MED (FR)  
 COMMENT  
 Other publication FR 2751000 19980116.  
 Location/Qualifiers  
 1. .286  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"  
 BASE COUNT 88 a 73 c 63 g 62 t  
 ORIGIN  
 Query Match 100.0%; Score 286; DB 6; Length 286;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-147;  
 Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AATTCGAGGAGCAGTACCGCAAGCGTTGCTCGCTATTCGCGGGTGATATAACAGAC 60  
 Db |||||  
 QY 1 AATTCGAGGAGCAGTACCGCAAGCGTTGCTCGCTATTCGCGGGTGATATAACAGAC 60  
 Db |||||  
 QY 61 GAGGGTATCCGCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 120  
 Db |||||  
 QY 61 GAGGGTATCCGCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 120  
 Db |||||  
 QY 121 AAAAAATGCTTTTTCATTTTTCGCAAGCAATGACGACAGCTCAGCCCAACCAACT 180  
 Db |||||  
 QY 121 AAAAAATGCTTTTTCATTTTTCGCAAGCAATGACGACAGCTCAGCCCAACCAACT 180  
 Db |||||  
 QY 181 GACCCATTGCGCATTTTATGAAAAGACGCTCAAAAAGGCATTATCAGATTGCGGCGT 240  
 Db |||||  
 QY 181 GACCCATTGCGCATTTTATGAAAAGACGCTCAAAAAGGCATTATCAGATTGCGGCGT 240  
 Db |||||  
 QY 241 AGACCCAGTGGAGAAAAGTTCAATGGCTCCAAACCATTCGCGAATT 286  
 Db |||||  
 QY 241 AGACCCAGTGGAGAAAAGTTCAATGGCTCCAAACCATTCGCGAATT 286  
 Db |||||

QY 1 AATTCGAGGAGCAGTACCGCAAGCGTTGCTCGCTATTCGCGGGTGATATAACAGAC 60  
 Db |||||  
 QY 1 AATTCGAGGAGCAGTACCGCAAGCGTTGCTCGCTATTCGCGGGTGATATAACAGAC 60  
 Db |||||  
 QY 61 GAGGGTATCCGCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 120  
 Db |||||  
 QY 61 GAGGGTATCCGCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 120  
 Db |||||  
 QY 121 AAAAAATGCTTTTTCATTTTTCGCAAGCAATGACGACAGCTCAGCCCAACCAACT 180  
 Db |||||  
 QY 121 AAAAAATGCTTTTTCATTTTTCGCAAGCAATGACGACAGCTCAGCCCAACCAACT 180  
 Db |||||  
 QY 181 GACCCATTGCGCATTTTATGAAAAGACGCTCAAAAAGGCATTATCAGATTGCGGCGT 240  
 Db |||||  
 QY 181 GACCCATTGCGCATTTTATGAAAAGACGCTCAAAAAGGCATTATCAGATTGCGGCGT 240  
 Db |||||  
 QY 241 AGACCCAGTGGAGAAAAGTTCAATGGCTCCAAACCATTCGCGAATT 286  
 Db |||||  
 QY 241 AGACCCAGTGGAGAAAAGTTCAATGGCTCCAAACCATTCGCGAATT 286  
 Db |||||  
 RESULT 3  
 BD063024  
 LOCUS  
 DEFINITION  
 286 bp DNA linear PAT 27-AUG-2002  
 DNA and specific proteins or peptides of the Neisseria meningitidis  
 species bacteria, method for obtaining them and their biological  
 applications.  
 ACCESSION  
 BD063024  
 VERSION  
 BD063024.1 GI:22608627  
 KEYWORDS  
 JP 2001504684-A/86.  
 SOURCE  
 unidentified  
 ORGANISM  
 unclassified.  
 REFERENCE  
 1 (bases 1 to 286)  
 Nassif,X., Tinsley,C., Achtman,M., Ruelle,J.L., Vinals,C. and  
 Merker,P.  
 TITLE  
 DNA and specific proteins or peptides of the Neisseria meningitidis  
 species bacteria, method for obtaining them and their biological  
 applications  
 JOURNAL  
 Patent: JP 2001504684-A 86 10-APR-2001;  
 INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, MAX  
 PLANCK GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV BERLIN,  
 SMITHKLINE BEECHAM  
 COMMENT  
 PN JP 2001504684-A/86  
 PD 10-APR-2001  
 PF 11-JUL-1997 JP 199805685  
 PR 12-JUL-1996 FR 96/08768  
 PI XAVIER NASSIF, COLIN TINSLEY, MARK ACHTMAN, JEAN LOUIS RUELLE, PI  
 CARLA VINALS,  
 PI PETRA MERKER  
 PC C12N15/31, C07K14/22, C07K16/12, A61K39/095, C12Q1/69, G01N33/53 CC  
 Strandedness: Single;  
 CC Topology: Linear;  
 FH Key Location/Qualifiers.  
 Location/Qualifiers  
 1. .286  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"  
 BASE COUNT 88 a 73 c 63 g 62 t  
 ORIGIN  
 Query Match 100.0%; Score 286; DB 6; Length 286;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-147;  
 Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AATTCGAGGAGCAGTACCGCAAGCGTTGCTCGCTATTCGCGGGTGATATAACAGAC 60  
 Db |||||  
 QY 1 AATTCGAGGAGCAGTACCGCAAGCGTTGCTCGCTATTCGCGGGTGATATAACAGAC 60  
 Db |||||  
 QY 61 GAGGGTATCCGCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 120  
 Db |||||  
 QY 61 GAGGGTATCCGCTGATGCAACAGAGCGATTACGGCAACTTGTCTTACCACATCCGTAAT 120  
 Db |||||



Query Match 61.9%; Score 177; DB 6; Length 3204;  
Best Local Similarity 100.0%; Pred. No. 9.7e-87;  
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 AATTCGAGGAGCAGTACCGCCAAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 60  
|||||  
Db 865 AATTCGAGGAGCAGTACCGCCAAAGCGTTGCTCGCCTATTCCGGCGGTGATAAAACAGAC 924  
|||||  
  
QY 61 GAGGCTATCGCCTGATGCACAGAGGATTACGGCAACTGTCTTACCACATCCGTAAT 120  
|||||  
Db 925 GAGGCTATCGCCTGATGCACAGAGGATTACGGCAACTGTCTTACCACATCCGTAAT 984  
|||||  
  
QY 121 AAAAAATGCTTTTTCATTTTTCGCGAAGCAATGACGACCAAGCTCAGCCCAACACA 177  
|||||  
Db 985 AAAAAATGCTTTTTCATTTTTCGCGAAGCAATGACGACCAAGCTCAGCCCAACACA 1041  
|||||

RESULT 6  
NMA222491/c 349061 bp DNA linear BCT 02-SEP-2002  
LOCUS Neisseria meningitidis serogroup A strain Z2491 complete genome;  
DEFINITION segment 2/7.  
ACCESSION AL162753 AL157959  
VERSION AL162753.2 GI:7379120  
KEYWORDS  
SOURCE Neisseria meningitidis Z2491  
ORGANISM Neisseria meningitidis Z2491  
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
Neisseriaceae; Neisseria.  
REFERENCE 1 (bases 1 to 349061)  
AUTHORS Parkhill, J., Achtman, M., James, K.D., Bentley, S.D., Churcher, C.,  
Klee, S.R., Morelli, G., Basham, D., Brown, D., Chillingworth, T.,  
Davies, R.M., Davis, P., Dellin, K., Feltwell, T., Hamlin, N.,  
Holroyd, S., Jagsels, K., Leather, S., Mungall, K.,  
Quail, M.A., Rajandream, M.A., Rutherford, K.M., Simmonds, M.,  
Skelton, J., Whitehead, S., Spratt, B.G. and Barrall, B.G.  
TITLE Complete DNA sequence of a serogroup A strain of Neisseria  
meningitidis Z2491  
NATURE 404 (6777), 502-506 (2000)  
JOURNAL 20222556  
MEDLINE 10761919  
PUBMED  
REFERENCE 2 (bases 1 to 349061)  
AUTHORS Parkhill, J.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAR-2000) Submitted on behalf of the Neisseria  
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,  
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk  
COMMENTS  
Notes:  
Details of N. meningitidis sequencing at the Sanger Centre are  
available on the World Wide Web.  
(URL, [http://www.sanger.ac.uk/Projects/N\\_meningitidis/](http://www.sanger.ac.uk/Projects/N_meningitidis/)).

FEATURES  
source  
1..349061  
/organism="Neisseria meningitidis Z2491"  
/mol\_type="genomic DNA"  
/strain="Z2491"  
/db\_xref="taxon:122587"  
/note="serogroup: A"  
complement (24..206)  
/note="ATP repeat; hmms hit to HMM ATR (1 - 183), score:  
310.39"  
/label=ATR  
RBS 209..212  
Gene 220..681  
CDS 220..681  
/gene="NMA0368"  
/note="NMA0368", probable integral membrane protein, len:  
133 aa; contains four probable transmembrane domains"  
/codon\_start=1  
/transl\_table=1  
/product="putative integral membrane protein"

/protein\_id="CAB83669.1"  
/db\_xref="GI:7379121"  
/db\_xref="SPTREMBL:Q9JWH7"  
/translation="MQEQRKSSPPIVMLVSLVIAWIASLGNVAFYLGHNHSGMEGLTV  
LILGIFASLDIRYCAVYVWLAAILVLLARKKVVFVHAAFWGLALVAFSVKAVYV  
DEAGNTSDIVRYGAGFYLVAAFAVASITGTFAGKNKERKAAASADGTNDV"  
671..1492  
/gene="hemk"  
671..1492  
/gene="hemk"  
/note="NMA0369, hemk, Hemk protein, len: 273 aa; similar  
to e.g. HEMK ECOLI P37186 HEMK protein (277 aa), fasta  
scores; E(): 0.42.3% identity in 279 aa overlap. Contains  
PS00092 N-6 Adenine-specific DNA methylases signature"  
/codon\_start=1  
/transl\_table=1  
/product="Hemk protein"  
/protein\_id="CAB83670.1"  
/db\_xref="GI:7379122"  
/db\_xref="SPTREMBL:Q9JWH6"  
/translation="WTFDKWLGLSKLPKNEARMMLQYSEYTRVOLLTRGGEMDPDEV  
RQRDLRAQRRLNGEVPVAYILGAREFFYGRRTVPNSVLIPRPETHLVEAVLARLPEN  
GRVMDLGTSGAVATVALERPDARFASDIPSPALETARKNAADLGARVEFYGSWF  
DTMPSEKGDIIIVSNPPYIENGDKHLSQGLRPEQIALTFSDGLSCIRTLAOGAP  
DRLAEGGFLLEHGFDDQGAAYRGLAENGFSVETLPDLAGLDRVTLGKYMKHLK"  
831..840  
/gene="hemk"  
/note="Core DNA uptake sequence: gccgtctgaa"  
/label=DUS  
1162..1171  
/gene="hemk"  
/note="Core DNA uptake sequence: gccgtctgaa"  
/label=DUS  
1187..1207  
/gene="hemk"  
/note="PS00092 N-6 Adenine-specific DNA methylases  
signature"  
complement (1279..1288)  
/note="Core DNA uptake sequence: gccgtctgaa"  
/label=DUS  
1565..1568  
1578..2966  
/gene="NMA0370"  
1578..2966  
/gene="NMA0370"  
/note="NMA0370, probable integral membrane protein, len:  
462 aa; similar to hypothetical proteins e.g. Y325\_HAEIN  
P44640 hypothetical protein HI0325 (450 aa), fasta scores;  
E(): 0.49.8% identity in 464 aa overlap"  
/codon\_start=1  
/transl\_table=1  
/product="putative integral membrane protein"  
/protein\_id="CAB83671.1"  
/db\_xref="GI:7379123"  
/db\_xref="SPTREMBL:Q9JWH5"  
/translation="MNAVVAVIVMLVLSRVHVLSLTIGAFVGGVAGMPLONIA  
DAQVSOAGILPVFNKLEGCAGKIALSYAMLGAFAMATISGLPQOLAGAVRKLNR  
GMEDSVRSRGEGVVKMLLSLILVGMMSQNIPIPIHTAFIMIVPEPLLVFNRLKIDR  
RLIACVITFTGLVTTMFLPYGFGAIFLNEILLGNHSAAPQDVKVNNVMAAFIPL  
GMLAGLLAFVYRKPRLYQSNADTAGNDAANRPOPSPAYSLAAVAIVCFAIQAL  
MYEDSLVGLMGLPFAVPMGLVINRDKANDVFGEIGKMAWVGFIMIAAQGFVAVNNA  
TGHQPLVSSMAIFGNSKGMALAMLVGLLVTMIGIGSSFSFTPLIAIVVPLCVGL  
GFSPLATVAIVGTAGALGDAGSPASDSTLGTMTGLNADGQHDHIRDSVIFTHINIP  
LLIAGWIAAMVLI"  
2967..3191  
/gene="slyx"  
2967..3191  
/gene="slyx"  
/note="NMA0371, slyx, SLXX protein homolog, len: 74 aa;  
similar to SLXX\_HAEIN P44759 SLXX protein homolog (73 aa),  
fasta scores; E(): 0.0017, 33.8% identity in 74 aa  
overlap, and SLXX\_ECOLI P30857 SLXX protein (72 aa), fasta  
scores; E(): 0.24, 32.4% identity in 68 aa overlap"



```
RNDISGTGLIKGGSQLQHGNNTYTKTIIIEGSLVLYGNKSDMRVETKGLIYN
GAASGSLNSDGIYVLADTDQSANETVHIKGSLOLDKGTLYTRGLKLLKVDSTAI
GKLYNSARGKAGYLNSTGRRVFLSAKIQDYSFFNIIETDGLLASLSDSEKTA
GSEGDTLSTVYRRGNAARTASAPAAAGLKHAVEGGSNLENLWELDASESATP
ETVETAAADRTDMPGIRPYGATFRAAAVQHANAADGVRIFNSLAATVYADSTAAD
MQORRLKAVSDGLDHNGTGLRIVTAQDQDGTWEGQGVGKMRGSTQTVGIAAKTGEN
TTAAATLGMGRSTWSENANAKTDSISLFPAGIRHDAGDIGYKGLFSYGRYKNSIRS
TGADEHAEGSVNGTLMOLGALGVNVPFAATGDLTVEGGLRYDILLKQDAPAFKGSALG
WGSNLSITEGTLVGLACLKLSQLSKAVLIFATAGVERDLNGRDYTVTGFTGATATG
KTGARMPHRLIVAGLIGADVEFGNGWNLARYSAGSKQYGNHSGRVGVGRNF"
79...81
/misc_feature
/genes="nalp"
/note="signal sequence cleavage site"
82..84
/misc_feature
/genes="nalp"
/note="lipid-binding site"
BASE COUNT      846 a   919 c   916 g   571 t
ORIGIN
Query Match      22.4%; Score 64; DB 1; Length 3252;
Best Local Similarity 100.0%; Pred. No. 5.7e-24;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 CTATTCCGGCGGTGATAAACACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGG 95
|||||
Db 906 CTATTCCGGCGGTGATAAACACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGG 965

QY 96 CAAC 99
|||||
Db 966 CAAC 969

RESULT 8
NME277537
LOCUS      3254 bp DNA linear BCT 24-JUL-2002
DEFINITION Neisseria meningitidis ausp gene for autotransporter serine
            protease.
ACCESSION  AJ277537
VERSION     AJ277537.2 GI:113539237
KEYWORDS   ausp gene; autotransporter serine protease.
SOURCE     Neisseria meningitidis
ORGANISM   Neisseria meningitidis
            Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
            Neisseriaceae; Neisseria.
REFERENCE  1
AUTHORS    Turner,D.P., Wooldridge,K.G. and Ala'Aldeen,D.A.
TITLE      Autotransported serine protease A of Neisseria meningitidis: an
            immunogenic, surface-exposed outer membrane, and secreted protein
JOURNAL    Infect. Immun. 70 (8), 4447-4461 (2002)
MEDLINE    22112893
PUBMED     12117956
REFERENCE  2
AUTHORS    Ala'Aldeen,D.A.A.
TITLE      Direct Submission
JOURNAL    Submitted (19-APR-2000) Ala'Aldeen D.A.A., Microbiology, University
            of Nottingham, Meningococcal Research Group, Division of
            Microbiology, University Hospital, Nottingham NG7 2UH, UNITED
            KINGDOM
REMARK     Revised by [3]
REFERENCE  3
AUTHORS    Ala'Aldeen,D.A.A.
TITLE      Direct Submission
JOURNAL    Submitted (02-APR-2001) Ala'Aldeen D.A.A., Microbiology, University
            of Nottingham, Meningococcal Research Group, Division of
            Microbiology, University Hospital, Nottingham NG7 2UH, UNITED
            KINGDOM
COMMENT    On Apr 3, 2001 this sequence version replaced gi:7649689.
FEATURES             Location/Qualifiers
     source           1..3254
                     /organism="Neisseria meningitidis"
                     /mol_type="genomic DNA"
                     /strain="B:15:P1.16"
                     /isolate="SD"

RNDISGTGLIKGGSQLQHGNNTYTKTIIIEGSLVLYGNKSDMRVETKGLIYN
GAASGSLNSDGIYVLADTDQSANETVHIKGSLOLDKGTLYTRGLKLLKVDSTAI
GKLYNSARGKAGYLNSTGRRVFLSAKIQDYSFFNIIETDGLLASLSDSEKTA
GSEGDTLSTVYRRGNAARTASAPAAAGLKHAVEGGSNLENLWELDASESATP
ETVETAAADRTDMPGIRPYGATFRAAAVQHANAADGVRIFNSLAATVYADSTAAD
MQORRLKAVSDGLDHNGTGLRIVTAQDQDGTWEGQGVGKMRGSTQTVGIAAKTGEN
TTAAATLGMGRSTWSENANAKTDSISLFPAGIRHDAGDIGYKGLFSYGRYKNSIRS
TGADEHAEGSVNGTLMOLGALGVNVPFAATGDLTVEGGLRYDILLKQDAPAFKGSALG
WGSNLSITEGTLVGLACLKLSQLSKAVLIFATAGVERDLNGRDYTVTGFTGATATG
KTGARMPHRLIVAGLIGADVEFGNGWNLARYSAGSKQYGNHSGRVGVGRNF"
79...81
/misc_feature
/genes="nalp"
/note="signal sequence cleavage site"
82..84
/misc_feature
/genes="nalp"
/note="lipid-binding site"
BASE COUNT      846 a   919 c   916 g   571 t
ORIGIN
Query Match      22.4%; Score 64; DB 1; Length 3252;
Best Local Similarity 100.0%; Pred. No. 5.7e-24;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 CTATTCCGGCGGTGATAAACACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGG 95
|||||
Db 906 CTATTCCGGCGGTGATAAACACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGG 965

QY 96 CAAC 99
|||||
Db 966 CAAC 969

RESULT 9
AX236409
LOCUS      3939 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 102 from Patent WO0164922.
ACCESSION  AX236409
VERSION     AX236409.1 GI:15796027
KEYWORDS   synthetic construct
            synthetic construct
            artificial sequences.
ORGANISM   Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
            Guillani,M. and Pizzi,M.C.
            Heterologous expression of neisserial proteins
            Patent: WO 0164922-A 102 07-SEP-2001;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source           1..3939
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="deltaG983-741"
BASE COUNT      1046 a  1100 c  1107 g   686 t
ORIGIN
Query Match      22.4%; Score 64; DB 6; Length 3939;
Best Local Similarity 100.0%; Pred. No. 5.6e-24;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 CTATTCCGGCGGTGATAAACACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGG 95
|||||
Db 801 CTATTCCGGCGGTGATAAACACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGG 860

QY 96 CAAC 99
|||||
Db 861 CAAC 864

RESULT 10
AX236419
LOCUS      3939 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 112 from Patent WO0164922.
ACCESSION  AX236419
VERSION     AX236419.1 GI:15796032
KEYWORDS
```

```

SOURCE      synthetic construct
ORGANISM     synthetic construct
KEYWORDS     artificial sequences.
REFERENCE    1
AUTHORS      Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
             Guillani,M.M. and Piza,M.C.
TITLE        Heterologous expression of neisserial proteins
JOURNAL      Patent: WO 0164922-A 112 07-SEP-2001;
             Chiron Spa (IT)
FEATURES     Location/Qualifiers
             source
               1..3939
               /organism="synthetic construct"
               /mol_type="genomic DNA"
               /db_xref="taxon:32630"
               /note="deltaG741-983"
BASE COUNT   1046 a 1103 c 1106 g 684 t
ORIGIN
Query Match 22.4%; Score 64; DB 6; Length 3939;
Best Local Similarity 100.0%; Pred. No. 5.6e-24;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY          36 CTATTCGGCGGTGATAAAACAGACGAGGGTATCCGCTGATGCAACAGAGCGATTACGG 95
             |||||||
Db          1569 CTATTCGGCGGTGATAAAACAGACGAGGGTATCCGCTGATGCAACAGAGCGATTACGG 1628
QY          96 CAAC 99
             |||||
Db          1629 CAAC 1632

RESULT 11
AX239717
LOCUS        AX239717
DEFINITION   Sequence 19 from Patent WO0164920.
ACCESSION    AX239717
VERSION      AX239717.1 GI:15797370
KEYWORDS     synthetic construct
SOURCE       synthetic construct
ORGANISM     artificial sequences.
REFERENCE    1
AUTHORS      Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Guillani,M.M.
             and Piza,M.
TITLE        Hybrid expression of neisserial proteins
JOURNAL      Patent: WO 0164920-A 19 07-SEP-2001;
             Chiron Spa (IT)
FEATURES     Location/Qualifiers
             source
               1..3939
               /organism="synthetic construct"
               /mol_type="genomic DNA"
               /db_xref="taxon:32630"
               /note="deltaG983-741"
BASE COUNT   1046 a 1100 c 1107 g 686 t
ORIGIN
Query Match 22.4%; Score 64; DB 6; Length 3939;
Best Local Similarity 100.0%; Pred. No. 5.6e-24;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY          36 CTATTCGGCGGTGATAAAACAGACGAGGGTATCCGCTGATGCAACAGAGCGATTACGG 95
             |||||||
Db          801 CTATTCGGCGGTGATAAAACAGACGAGGGTATCCGCTGATGCAACAGAGCGATTACGG 860
QY          96 CAAC 99
             |||||
Db          861 CAAC 864

RESULT 12
AX239729
LOCUS        AX239729
DEFINITION   Sequence 31 from Patent WO0164920.

```

```

ACCESSION    AX239729
VERSION      AX239729.1 GI:15797375
KEYWORDS     synthetic construct
SOURCE       synthetic construct
ORGANISM     artificial sequences.
REFERENCE    1
AUTHORS      Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Guillani,M.M.
             and Piza,M.
TITLE        Hybrid expression of neisserial proteins
JOURNAL      Patent: WO 0164920-A 31 07-SEP-2001;
             Chiron Spa (IT)
FEATURES     Location/Qualifiers
             source
               1..3939
               /organism="synthetic construct"
               /mol_type="genomic DNA"
               /db_xref="taxon:32630"
               /note="deltaG741-983"
BASE COUNT   1046 a 1103 c 1106 g 684 t
ORIGIN
Query Match 22.4%; Score 64; DB 6; Length 3939;
Best Local Similarity 100.0%; Pred. No. 5.6e-24;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY          36 CTATTCGGCGGTGATAAAACAGACGAGGGTATCCGCTGATGCAACAGAGCGATTACGG 95
             |||||||
Db          1569 CTATTCGGCGGTGATAAAACAGACGAGGGTATCCGCTGATGCAACAGAGCGATTACGG 1628
QY          96 CAAC 99
             |||||
Db          1629 CAAC 1632

RESULT 13
AX236463
LOCUS        AX236463
DEFINITION   Sequence 156 from Patent WO0164922.
ACCESSION    AX236463
VERSION      AX236463.1 GI:15796053
KEYWORDS     synthetic construct
SOURCE       synthetic construct
ORGANISM     artificial sequences.
REFERENCE    1
AUTHORS      Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
             Guillani,M.M. and Piza,M.C.
TITLE        Heterologous expression of neisserial proteins
JOURNAL      Patent: WO 0164922-A 156 07-SEP-2001;
             Chiron Spa (IT)
FEATURES     Location/Qualifiers
             source
               1..4170
               /organism="synthetic construct"
               /mol_type="genomic DNA"
               /db_xref="taxon:32630"
               /note="961C-983"
BASE COUNT   1185 a 1138 c 1119 g 728 t
ORIGIN
Query Match 22.4%; Score 64; DB 6; Length 4170;
Best Local Similarity 100.0%; Pred. No. 5.6e-24;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY          36 CTATTCGGCGGTGATAAAACAGACGAGGGTATCCGCTGATGCAACAGAGCGATTACGG 95
             |||||||
Db          1800 CTATTCGGCGGTGATAAAACAGACGAGGGTATCCGCTGATGCAACAGAGCGATTACGG 1859
QY          96 CAAC 99
             |||||
Db          1860 CAAC 1863

RESULT 14

```

Search completed: November 6, 2003, 08:21:46  
Job time : 1523 secs

AX239749  
LOCUS AX239749 4170 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 51 from Patent WO0164920.  
ACCESSION AX239749  
VERSION AX239749.1 GI:15797385  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Arico,M.B., Comanducci,M., Galeotti,C., Massignani,V., Giuliani,M.M.  
and Pizza,M.  
TITLE Hybrid expression of neisserial proteins  
JOURNAL Patent: WO 0164920-A 51 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
source  
1..4170  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="961c-983"  
BASE COUNT 1185 a 1138 c 1119 g 728 t  
ORIGIN

Query Match 22.4%; Score 64; DB 6; Length 4170;  
Best Local Similarity 100.0%; Pred. No. 5.6e-24;  
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 36 CTATTCCGGCGGTGATAAACAACACAGAGGGTATCCGCTGATGCAACACAGAGCGATTACGG 95  
Db 1800 CTATTCCGGCGGTGATAAACAACACAGAGGGTATCCGCTGATGCAACACAGAGCGATTACGG 1859

QY 96 CAAC 99  
Db 1860 CAAC 1863

RESULT 15  
AX236413  
LOCUS AX236413 4179 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 106 from Patent WO0164922.  
ACCESSION AX236413  
VERSION AX236413.1 GI:15796029  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C.C., Massignani,V.C.,  
Guiliani,M.M. and Pizza,M.C.  
TITLE Heterologous expression of neisserial proteins  
JOURNAL Patent: WO 0164922-A 106 07-SEP-2001;  
Chiron Spa (IT)  
FEATURES  
source  
1..4179  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="deltaG983-961c"  
BASE COUNT 1187 a 1140 c 1121 g 731 t  
ORIGIN

Query Match 22.4%; Score 64; DB 6; Length 4179;  
Best Local Similarity 100.0%; Pred. No. 5.6e-24;  
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 36 CTATTCCGGCGGTGATAAACAACACAGAGGGTATCCGCTGATGCAACACAGAGCGATTACGG 95  
Db 801 CTATTCCGGCGGTGATAAACAACACAGAGGGTATCCGCTGATGCAACACAGAGCGATTACGG 860

QY 96 CAAC 99  
Db 861 CAAC 864



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2003, 08:14:50 ; Search time 204 Seconds  
(without alignments)  
4473.598 Million cell updates/sec

Title: US-09-928-457-95  
Perfect score: 286  
Sequence: 1 AATTCGAGGAGCAGTACCG.....GCTCAACCATTCGGAAATT 286

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2141354 seqs, 1595478879 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 4282708

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	286	100.0	286	10	US-09-928-457-95
2	18	6.3	655	12	US-10-027-632-206121
3	18	6.3	655	13	US-10-027-632-206121
4	17	5.9	257	8	US-08-834-666A-47
5	17	5.9	308	8	US-08-834-666A-44
6	17	5.9	490	10	US-09-881-752A-225
7	17	5.9	534	12	US-10-027-632-223041
8	17	5.9	534	13	US-10-027-632-223041
9	17	5.9	540	12	US-09-882-227-319
10	17	5.9	547	12	US-10-027-632-43992
11	17	5.9	547	13	US-10-027-632-43992
12	17	5.9	596	12	US-10-027-632-253523
13	17	5.9	596	13	US-10-027-632-253523
14	17	5.9	623	12	US-10-027-632-75473
15	17	5.9	623	13	US-10-027-632-75473
16	17	5.9	624	12	US-10-027-632-45441

c	17	5.9	624	13	US-10-027-632-45441	Sequence 45441, A
c	18	5.9	652	12	US-10-027-632-59744	Sequence 59744, A
c	19	5.9	652	13	US-10-027-632-59744	Sequence 59744, A
c	20	5.9	663	12	US-10-027-632-224318	Sequence 224318, A
c	21	5.9	663	13	US-10-027-632-224318	Sequence 224318, A
c	22	5.9	741	12	US-10-027-632-128961	Sequence 128961, A
c	23	5.9	741	13	US-10-027-632-128961	Sequence 128961, A
c	24	5.9	914	12	US-10-027-632-121898	Sequence 121898, A
c	25	5.9	914	13	US-10-027-632-121898	Sequence 121898, A
c	26	5.9	914	13	US-10-027-632-121898	Sequence 121898, A
c	27	5.9	914	13	US-10-027-632-121898	Sequence 121898, A
c	28	5.9	1131	12	US-10-027-632-117943	Sequence 117943, A
c	29	5.9	1131	13	US-10-027-632-117943	Sequence 117943, A
c	30	5.9	1968	12	US-10-349-680-25	Sequence 25, Appl
c	31	5.9	2161	8	US-08-834-666A-17	Sequence 17, Appl
c	32	5.9	2248	8	US-08-834-666A-15	Sequence 15, Appl
c	33	5.9	9990	9	US-09-767-088A-2	Sequence 2, Appl
c	34	5.9	9990	9	US-09-767-088A-15	Sequence 15, Appl
c	35	5.9	368004	10	US-09-949-654-3	Sequence 3, Appl
c	36	5.9	1830121	14	US-10-329-960-1	Sequence 1, Appl
c	37	5.6	142	9	US-09-770-696-549	Sequence 549, App
c	38	5.6	256	13	US-10-040-739-812	Sequence 812, App
c	39	5.6	301	10	US-09-974-300-6395	Sequence 6395, Ap
c	40	5.6	348	12	US-10-091-007-167	Sequence 167, App
c	41	5.6	352	10	US-09-960-352-6025	Sequence 6025, Ap
c	42	5.6	546	12	US-10-027-632-203483	Sequence 203483, A
c	43	5.6	546	12	US-10-027-632-203484	Sequence 203484, A
c	44	5.6	546	13	US-10-027-632-203483	Sequence 203483, A
c	45	5.6	546	13	US-10-027-632-203484	Sequence 203484, A

ALIGNMENTS

RESULT 1

US-09-928-457-95  
; Sequence 95, Application US/09928457  
; Patent No. US20020164603A1  
; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: DNA, specific proteins and peptides  
; TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method  
; TITLE OF INVENTION: for obtaining them and their biological application.

; NUMBER OF SEQUENCES: 99

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30 (ORB)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/928,457

; FILING DATE: 2001-08-14

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 09/214,759

; FILING DATE: 199-12-10

; INFORMATION FOR SEQ ID NO: 95:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 286 base pairs

; TYPE: nucleotide

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHEetical: NO

; ANTI-SENSE: NO

US-09-928-457-95

Query Match 100.0%; Score 286; DB 10; Length 286;

Best Local Similarity 100.0%; Pred. No. 3.2e-148;

Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTCGAGGAGCAGTACCGCCCAAGCGTTCCTCGCCTATTTCGGCGGTGATAAACAGAC 60

Db 1 AATTCGAGGAGCAGTACCGCCCAAGCGTTCCTCGCCTATTTCGGCGGTGATAAACAGAC 60

```

QY 61 GAGGGTATCGGCTGATGCAACAGAGCGGATTACGGCAACTTGTCTTACCACATCCGTAAT 120
DB 61 GAGGGTATCGGCTGATGCAACAGAGCGGATTACGGCAACTTGTCTTACCACATCCGTAAT 120
QY 121 AAAACATGCTTTTTCATTTTTCGGCAGCAATGACGCAAGCTCAGCCCAACACAACT 180
DB 121 AAAACATGCTTTTTCATTTTTCGGCAGCAATGACGCAAGCTCAGCCCAACACAACT 180
QY 181 GACCCCTATTGCCATTATGAAGAGAGCGCTCAAAAGGCATTATCACAGTTGCGAGGCT 240
DB 181 GACCCCTATTGCCATTATGAAGAGAGCGCTCAAAAGGCATTATCACAGTTGCGAGGCT 240
QY 241 AGACCCGACGTCGAGAGAAAGCTTCAATGGCTCCCAACCATTCGCGAAT 286
DB 241 AGACCCGACGTCGAGAGAAAGCTTCAATGGCTCCCAACCATTCGCGAAT 286

```

```

RESULT 2
US-10-027-632-206121/c
; Sequence 206121, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 206121
; LENGTH: 655
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(655)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-206121

```

```

Query Match 6.3%; Score 18; DB 12; Length 655;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 116 GTAATAAAACATGCTTT 133
DB 572 GTAATAAAACATGCTTT 555

```

```

RESULT 3
US-10-027-632-206121/c
; Sequence 206121, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632

```

```

; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 206121
; LENGTH: 655
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(655)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-206121

```

```

Query Match 6.3%; Score 18; DB 13; Length 655;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 116 GTAATAAAACATGCTTT 133
DB 572 GTAATAAAACATGCTTT 555

```

```

RESULT 4
US-08-834-666A-47
; Sequence 47, Application US/08834666A
; Publication No. US20020044949A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Lissolo, Ling
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Miller, Charles
; APPLICANT: Al-Garawi, Amal
; TITLE OF INVENTION: 76 kDa Helicobacter Polypeptides and
; Corresponding Polynucleotide Molecules
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,666A
; FILING DATE: 01-APR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,175
; REFERENCE/DOCKET NUMBER: 06132/038001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:

```

; INFORMATION FOR SEQ ID NO: 47:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 29 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: Genomic DNA

US-08-834-666A-47

Query Match 5.9%; Score 17; DB 8; Length 29;

Best Local Similarity 100.0%; Pred. No. 49;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 192 CATTATTGAAAAGAC 208

Db 11 CATTATTGAAAAGAC 27

RESULT 5

US-08-834-666A-44

; Sequence 44, Application US/08834666A

; Publication No. US200200449A1

; GENERAL INFORMATION:

; APPLICANT: Kleanthous, Harold

; APPLICANT: Lissolo, Ling

; APPLICANT: Tomb, Jean-Francois

; APPLICANT: Miller, Charles

; APPLICANT: Al-Garawi, Amal

; TITLE OF INVENTION: 76 kDa Helicobacter Polypeptides and

; Corresponding Polynucleotide Molecules

; NUMBER OF SEQUENCES: 64

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Clark & Elbing LLP

; STREET: 176 Federal Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/834,666A

; FILING DATE: 01-APR-1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Paul T.

; REGISTRATION NUMBER: 30,175

; REFERENCE/DOCKET NUMBER: 06132/038001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-428-0200

; TELEFAX: 617-428-7045

; TELEX:

; INFORMATION FOR SEQ ID NO: 44:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 30 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: Genomic DNA

US-08-834-666A-44

Query Match

Best Local Similarity 5.9%; Score 17; DB 8; Length 30;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 192 CATTATTGAAAAGAC 208

Db 11 CATTATTGAAAAGAC 27

RESULT 6

US-09-881-752A-225/c

; Sequence 225, Application US/09881752A

; Patent No. US20020115078A1

; GENERAL INFORMATION:

; APPLICANT: Kleanthous, Harold

; APPLICANT: Al-Garawi, Amal

; APPLICANT: Miller, Charles

; APPLICANT: Tomb, Jean-Francois

; APPLICANT: Omen, Raymond P.

; TITLE OF INVENTION: Identification of polynucleotides

; Encoding No. US20020115078A1 Helicobacter Polypeptides in the

; TITLE OF INVENTION: Genome

; FILE REFERENCE: 06132/041002

; CURRENT APPLICATION NUMBER: US/09/881,752A

; CURRENT FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: US 08/833,457

; PRIOR FILING DATE: 1997-04-01

; NUMBER OF SEQ ID NOS: 370

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 225

; LENGTH: 490

; TYPE: DNA

; ORGANISM: Helicobacter pylori

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (51)...(437)

US-09-881-752A-225

Query Match

Best Local Similarity 5.9%; Score 17; DB 10; Length 490;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 ATGCTTTTCATTTTTC 143

Db 61 ATGCTTTTCATTTTTC 45

RESULT 7

US-10-027-632-229041/c

; Sequence 229041, Application US/10027632

; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 229041

; LENGTH: 534

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)...(534)

; OTHER INFORMATION: n = A,T,C or G

```
US-10-027-632-229041
Query Match          5.9%; Score 17; DB 12; Length 534;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 192 CATTATGAAAGAC 208
    |||||
Db 202 CATTATGAAAGAC 186

RESULT 8
US-10-027-632-229041/c
; Sequence 319, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 229041
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..:(534)
; OTHER INFORMATION: n = A, T, C or G
US-10-027-632-229041

Query Match          5.9%; Score 17; DB 13; Length 534;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 192 CATTATGAAAGAC 208
    |||||
Db 202 CATTATGAAAGAC 186

RESULT 9
US-09-882-227-319/c
; Sequence 319, Application US/09882227
; Publication No. US20030158396A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20030158396A1 Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/047002
; CURRENT APPLICATION NUMBER: US/09/882,227
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/902,615
```

```
US-09-882-227-319
; PRIOR FILING DATE: 1997-07-29
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 319
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (16)...(513)
US-09-882-227-319

Query Match          5.9%; Score 17; DB 12; Length 540;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 ATGCTTTTCATTTTTC 143
    |||||
Db 523 ATGCTTTTCATTTTTC 507

RESULT 10
US-10-027-632-43992
; Sequence 43992, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43992
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-43992

Query Match          5.9%; Score 17; DB 12; Length 547;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 AATAAAACATGCTTTT 134
    |||||
Db 324 AATAAAACATGCTTTT 340

RESULT 11
US-10-027-632-43992
; Sequence 43992, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
```

; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 43992  
; LENGTH: 547  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-43992

Query Match 5.9%; Score 17; DB 13; Length 547;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 AATAAAACATGCTTTT 134  
|||||  
Db 324 AATAAAACATGCTTTT 340

## RESULT 12

US-10-027-632-253523  
; Sequence 253523, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 253523  
; LENGTH: 596  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-253523

Query Match 5.9%; Score 17; DB 12; Length 596;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 AATAAAACATGCTTTT 134  
|||||  
Db 283 AATAAAACATGCTTTT 299

## RESULT 13

US-10-027-632-253523  
; Sequence 253523, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 253523  
; LENGTH: 596  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-253523

Query Match 5.9%; Score 17; DB 13; Length 596;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 AATAAAACATGCTTTT 134  
|||||  
Db 283 AATAAAACATGCTTTT 299

## RESULT 14

US-10-027-632-75473  
; Sequence 75473, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 75473  
; LENGTH: 623  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-75473



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2003, 07:39:22 ; Search time 1332 Seconds  
(without alignments)  
5218.529 Million cell updates/sec

Title: US-09-928-457-95

Perfect score: 286  
Sequence: 1 AATTCGAGAGAGCAGTACCG.....GCTCCAAACATTCGCGAATT 286

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 1215238056 residues

Word size: 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rnd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	22	7.7	947	28	AQ747328 HS_5536 A
C 2	21	7.3	803	29	AL331132 Tetradon
C 3	21	7.3	1068	29	CNS05RXP
C 4	20	7.0	244	12	BI321685

5	20	7.0	375	28	AZ470268
C 6	20	7.0	646	13	BU362895
C 7	20	7.0	734	13	BU134402
C 8	20	7.0	824	13	BU421963
C 9	19	6.6	253	29	CC104967
C 10	19	6.6	286	9	AL838311
C 11	19	6.6	416	14	CB767580
C 12	19	6.6	609	29	AG247998
C 13	19	6.6	629	9	AV649619
C 14	19	6.6	646	14	CA360393
C 15	19	6.6	655	28	BH374446
C 16	19	6.6	677	14	CB945056
C 17	19	6.6	738	14	CB461162
C 18	19	6.6	766	29	BZ774016
C 19	19	6.6	814	19	CB680563
C 20	19	6.6	836	13	BU219080
C 21	19	6.6	927	14	CA985703
C 22	19	6.6	941	10	BE783156
C 23	18	6.3	118	9	AW708999
C 24	18	6.3	185	10	BF818300
C 25	18	6.3	190	9	AW721882
C 26	18	6.3	192	9	AW711311
C 27	18	6.3	207	9	AI319959
C 28	18	6.3	246	9	AI319735
C 29	18	6.3	254	9	AW722970
C 30	18	6.3	258	10	BF290318
C 31	18	6.3	263	9	AW709370
C 32	18	6.3	280	9	AI176829
C 33	18	6.3	284	9	AW718106
C 34	18	6.3	294	9	AI320261
C 35	18	6.3	294	9	AL918457
C 36	18	6.3	314	9	AW718452
C 37	18	6.3	326	9	AW711619
C 38	18	6.3	347	10	BF290317
C 39	18	6.3	361	10	BF401635
C 40	18	6.3	366	9	AW715053
C 41	18	6.3	366	9	AW722189
C 42	18	6.3	366	9	AW722424
C 43	18	6.3	366	9	AW722437
C 44	18	6.3	368	9	AW722960
C 45	18	6.3	375	9	AI320908

## ALIGNMENTS

RESULT 1  
AQ747328/c  
LOCUS  
DEFINITION HS\_5536\_A2\_C06\_T7A\_RPCT-11 Human Male BAC Library Homo sapiens  
genomic Clone Plate=1112 Col=12 Row=E, genomic survey sequence.  
ACCESSION AQ747328  
VERSION AQ747328.1 GI:5534486  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887

Email: jwallace@u.washington.edu  
 Clones are derived from the human BAC library RPCL-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACpac Resources (http://bacpac.med.buffalo.edu/ordering\_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.hbrc.washington.edu

Plate: 1112 row: E column: 12  
 Seq primer: T7  
 Class: BAC ends  
 High quality sequence stop: 947.  
 Location/Qualifiers

1.947

FEATURES  
 source  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate=1112 Col=12 Row=E"  
 /sex="male"  
 /clone\_lib="RPCL-11 Human Male BAC Library"  
 /note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

BASE COUNT  
 298 a 167 c 169 g 311 t 2 others

ORIGIN

Query Match 7.7%; Score 22; DB 28; Length 947;  
 Best Local Similarity 100.0%; Pred. No. 0.62;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ATTTATGAAAGACGCTCA 214  
 |||||  
 Db 585 ATTTATGAAAGACGCTCA 564

RESULT 2  
 CENS05E0W 803 bp DNA linear GSS 01-SEP-2000  
 LOCUS  
 DEFINITION  
 Tetracodon nigroviridis genome survey sequence T7 end of clone 009N18 of library B from Tetracodon nigroviridis, genomic survey sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 AL333113.1 GI:8226871  
 GSS: genome survey sequence.  
 Tetracodon nigroviridis  
 Tetracodon nigroviridis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.

REFERENCE

AUTHORS

1 Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fzames,C., Wincker,P., Brotlier,P., Quetier,F., Saurin,W. and Weissenbach,J.  
 Estimate of human gene number provided by genome-wide analysis using Tetracodon nigroviridis DNA sequence

JOURNAL  
 MEDLINE  
 PUBMED  
 20296633  
 10835645

REFERENCE

AUTHORS

2 Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C., Fzames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.  
 Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetracodon nigroviridis

JOURNAL  
 MEDLINE  
 PUBMED  
 20359837  
 10899143

REFERENCE

AUTHORS

3 (bases 1 to 803)  
 Genoscope.  
 Direct Submission  
 Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetracodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetracodon.

FEATURES

source

1.803  
 /organism="Tetracodon nigroviridis"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:99883"  
 /clone="009N18"  
 /clone\_lib="B"  
 /note="Genoscope sequence ID : COAB009D09C1-end : T7"

BASE COUNT  
 180 a 167 c 200 g 223 t 33 others

ORIGIN

Query Match 7.3%; Score 21; DB 29; Length 803;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 AAAACATGCTTTTCATTTT 141  
 |||||  
 Db 377 AAAACATGCTTTTCATTTT 397

RESULT 3

CNS05RXF/c

LOCUS

DEFINITION

Tetracodon nigroviridis genome survey sequence Sp6 end of clone 042M10 of library B from Tetracodon nigroviridis, genomic survey sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 AL351132.1 GI:8244902  
 GSS: genome survey sequence.  
 Tetracodon nigroviridis  
 Tetracodon nigroviridis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.

REFERENCE

AUTHORS

1 Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fzames,C., Wincker,P., Brotlier,P., Quetier,F., Saurin,W. and Weissenbach,J.  
 Estimate of human gene number provided by genome-wide analysis using Tetracodon nigroviridis DNA sequence

JOURNAL  
 MEDLINE  
 PUBMED  
 20296633  
 10835645

REFERENCE

AUTHORS

2 Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C., Fzames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.  
 Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetracodon nigroviridis

JOURNAL  
 MEDLINE  
 PUBMED  
 20359837  
 10899143

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetracodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetracodon.

FEATURES

source

1.1068  
 /organism="Tetracodon nigroviridis"  
 /mol\_type="genomic DNA"



/db\_xref="taxon:99883"  
/clone="042M10"  
/clone\_lib="B"  
/note="Genoscope sequence ID : COAR042BG05BI-end : SP6"

BASE COUNT 289 a 252 c 236 g 270 t 21 others

ORIGIN

Query Match 7.3%; Score 21; DB 29; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 AAAACATGCTTTTCAATTTT 141  
|||||  
776 AAAACATGCTTTTCAATTTT 756

Db

RESULT 4  
BI321685/c 244 bp mRNA linear EST 29-NOV-2001

LOCUS sat91d05.v1 Gm-c1079 Glycine max cDNA clone GENOME SYSTEMS CLONE

DEFINITION ID: Gm-c1079-1306.5' similar to SW:PPCK CUCSA P42066  
PHOSPHENOLPYRUVATE CARBOXYKINASE [ATP] ;, mRNA sequence.

ACCESSION BI321685 GI:15000871

VERSION BI321685.1 GI:15000871

KEYWORDS EST.

SOURCE Glycine max (soybean)

ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 244)  
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna  
A., Bolla, B., Maria, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wylie, T., Underwood, K., Steptoe, W., Theising, B., Allen, M., Bowers  
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harey, N., Bowers  
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished  
Contact: Shoemaker R./Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Trice considered overall poor quality This clone is available  
through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway  
Huntsville, AL 35801 For further information call: (800)-533-4363  
or contact via email: cnu@resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 1.  
Location/Qualifiers  
1..244  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1079-1306"  
/tissue\_type="Roots of 8 day old 'Brags' seedlings"  
/dev\_stage="8 days old"  
/lab\_host="MDH10P"  
/clone\_lib="Gm-c1079"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI; The mRNA was isolated from roots of 8 day old  
'Brags' seedlings that were mock-infected 72 hours prior  
to harvest. Dr. Gary Stacey generously donated the  
tissue. The roots were flash-frozen in liquid nitrogen.  
Stratagene's cDNA Synthesis Kit (catalog number 200401)  
was used to synthesize the cDNA. First-strand synthesis  
was performed with 5-methyl dCTP, hence the ligated cDNA  
was hemimethylated. A modification of Stratagene's  
first-strand synthesis primer was used. An 'anchor'

FEATURES  
Source

nucleotide (V=A, C, or G) was added to the 3' end of the  
primer [GAGAGAGAGAGAGAGAGACAGCTCGAG(T)18V] to anchor  
the primer at the 5' end of the poly(A) tract. After  
second-strand synthesis, the cDNA ends were filled in with  
cloned Pfu DNA, ligated to EcoRI adapters and subsequently  
phosphorylated. The cDNA was then precipitated and  
redissolved in sterile, RNase-, DNase-free water. The XhoI  
site within the first-strand synthesis primer was then  
restricted by digestion with XhoI from Promega (40U/ul);  
all XhoI sites in the cDNA would be protected by their  
hemimethylated status. The cDNA constructs were  
size-fractionated with a 500p cutoff, using Sephacryl  
S-500 High Resolution (Pharmacia Biotech) in a 2-mm  
diameter column and a bed volume of approximately 1ml. The  
column eluent was precipitated, redissolved, and ligated  
into Stratagene's pBluescript II XR Predigested vector  
(pBluescript II SK(+)) vector that has been digested with  
EcoRI and XhoI, and phosphorylated by Stratagene). This  
library was constructed in the laboratory of Dr. Paul Keim  
and Dr. Virginia H. Coryell at Northern Arizona  
University."

BASE COUNT 70 a 49 c 53 g 71 t 1 others

ORIGIN

Query Match 7.0%; Score 20; DB 12; Length 244;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 124 AACATGCTTTTCAATTTTTC 143  
|||||  
238 AACATGCTTTTCAATTTTTC 219

Db

RESULT 5  
A2470268 375 bp DNA linear GSS 04-OCT-2000

LOCUS 1M0284HI3F Mouse 10kb plasmid UNGCM library Mus musculus genomic

DEFINITION clone UNGCM0284HI3 F, genomic survey sequence.

ACCESSION A2470268 GI:10628393

VERSION A2470268.1 GI:10628393

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 375)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly  
'M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.  
and Wright, D., Weiss, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0284 row: H column: 13  
Seq primer: CGTTGTAAACGACGCGCACT  
Class: plasmid ends  
High quality sequence stop: 375.  
Location/Qualifiers  
1..375  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UNGCM0284HI3"

FEATURES  
Source

```

/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid U06C1M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g1|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT      100 a      80 c      78 g      117 t
ORIGIN

Query Match      7.0%; Score 20; DB 28; Length 375;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 20; Conservative 0; Mismatches 0; Gaps 0;

Cy      115 CGTATATAAACATGCTTTT 134
      |||||
Db      142 CGTATATAAACATGCTTTT 161

RESULT 6
BU362895/c      646 bp      mRNA      linear      EST 28-NOV-2002
LOCUS      603788707F1 CSEQCHN72 Gallus gallus cDNA clone CHEST748n1 5', mRNA
DEFINITION      sequence.
ACCESSION      BU362895
VERSION      BU362895.1 GI:25870896
KEYWORDS      EST.
SOURCE      Gallus gallus (chicken)
ORGANISM      Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 646)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
12435392
JOURNAL MEDLINE
PUBMED
COMMENT      Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .646
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton line 151"
/db_xref="taxon:9031"
/clone="CHEST748n1"
/sex="Female"
/tissue_type="cerebrum"
/dev_stage="adult"
/lab_host="DH10B"

```

```

/clone_lib="CSEQCHN72"
/notes="Organ: brain; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adaptors, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

BASE COUNT      217 a      139 c      118 g      172 t
ORIGIN

Query Match      7.0%; Score 20; DB 13; Length 646;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 20; Conservative 0; Mismatches 0; Gaps 0;

Cy      123 AACATGCTTTTCATTTT 142
      |||||
Db      155 AACATGCTTTTCATTTT 136

RESULT 7
BU134402/c      734 bp      mRNA      linear      EST 25-NOV-2002
LOCUS      603119331F1 CSEQCHL22 Gallus gallus cDNA clone CHEST78n8 5', mRNA
DEFINITION      sequence.
ACCESSION      BU134402
VERSION      BU134402.1 GI:25347142
KEYWORDS      EST.
SOURCE      Gallus gallus (chicken)
ORGANISM      Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 734)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
12435392
JOURNAL MEDLINE
PUBMED
COMMENT      Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .734
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST78n8"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCHL22"
/notes="Organ: heads; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; Modification of pBluescript II KS(+)
[Stratagene] vector to accommodate cDNA produced with the
T-trimmed protocol (Construction of uni-directionally
cloned cDNA libraries from messenger RNA for improved 3'
end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387
,624). Cut pBluescript II KS(+) with NotI and EcoRI.
Ligate in double stranded adaptor containing BspI and

```

BamHI sites [5'ggcgcgctgcagcccgatccgaaaaaaag]  
[5'aattcttttcgagatccggcgacgcgc]

BASE COUNT  
ORIGIN

251 a 134 c 161 g 188 t

Query Match 7.0%; Score 20; DB 13; Length 724;  
Best Local Similarity 100.0%; Pred. No. 8.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 123 AACATGCTTTTCATTTTT 142  
|||||  
Db 528 AACATGCTTTTCATTTTT 509

RESULT 8

BU421963/c 824 bp mRNA linear EST 29-NOV-2002  
LOCUS 603954560F1 CSEORBN09 Gallus gallus CDNA clone CHEST920g15 5', mRNA  
DEFINITION sequence.

ACCESSION BU421963  
VERSION BU421963.1 GI:25914634  
KEYWORDS EST.  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 824)  
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,  
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.  
A Comprehensive Collection of Chicken cDNAs  
Curr. Biol. 12 (22), 1965-1969 (2002)  
22335534  
12445392  
JOURNALS MEDLINE  
PUBMED

REFERENCE

AUTHORS

Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,  
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.  
A Comprehensive Collection of Chicken cDNAs  
Curr. Biol. 12 (22), 1965-1969 (2002)  
22335534  
12445392

COMMENT

Contact: Simon Hubbard  
Department of Biomolecular Sciences  
University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK  
Tel: 01612008930  
Fax: 01612360409  
Email: Simon.Hubbard@umist.ac.uk.

location/Qualifiers

1..824

/organism="Gallus gallus"

/mol\_type="mRNA"

/strain="Layer and broiler"

/db\_xref="taxon:9031"

/clone="CHEST920g15"

/sex="Male and female"

/tissue\_type="Chondrocytes isolated from growth plate  
cartilage"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="CSEORBN09"

/note="Vector: pBluescript II KS(+); Site\_1: EcoRI;  
Site\_2: NotI; This normalized library was constructed from  
1 million independent clones. CDNA synthesis was initiated  
using an oligo(dT) primer, using methylated C in the first  
strand synthesis reaction. Following this first strand  
reaction, double-stranded cDNA was blunted, ligated to  
NotI adapters, digested with EcoRI, size-selected, and  
cloned into the NotI and EcoRI compatible sites of a  
custom modified MCS of the pBluescript (KS+) vector. The  
library was normalized in 2 rounds using conditions  
adapted from Soares et al., PNAS (1994) 91: 9228-9232 and  
Bonaldi et al., Genome Research 6 (1996): 791, except that  
a significantly longer reannealing hybridization was  
used."

BASE COUNT 277 a 175 c 140 g 232 t

Query Match

7.0%; Score 20; DB 13; Length 824;

Best Local Similarity 100.0%; Pred. No. 8.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 123 AACATGCTTTTCATTTTT 142  
|||||  
Db 124 AACATGCTTTTCATTTTT 105

RESULT 9

CC104967/c 253 bp DNA linear GSS 16-APR-2003  
LOCUS CSU-K34.131K12.SP6 CSU-K34 Aedes aegypti genomic clone  
DEFINITION CSU-K34.131K12, genomic survey sequence.

ACCESSION CC104967  
VERSION CC104967.1 GI:29974023  
KEYWORDS GSS.  
SOURCE Aedes aegypti (yellow fever mosquito)  
ORGANISM Aedes aegypti  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.  
1 (bases 1 to 253)  
Loftus,B., Shetty,J., Severson,D., Brown,S. and Knudson,D.  
End sequencing of Aedes aegypti BACs  
Unpublished  
Other GSSs: CSU-K34.131K12.T7  
Contact: Brendan Loftus  
Department of Eukaryotic Genomics  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-3543  
Fax: 301-838-0208  
Email: entae@tigr.org  
Library was provided by Susan Brown and Dennis Knudson at Colorado  
State University.  
Seq primer: SP6  
Class: BAC ends.

location/Qualifiers

1..253

/organism="Aedes aegypti"

/mol\_type="genomic DNA"

/db\_xref="taxon:7159"

/clone="CSU-K34.131K12"

/clone\_lib="CSU-K34"

/note="Vector: pBAC3.6; Site\_1: EcoRI; Source DNA: Aedes  
aegypti; strain unknown (derived from freshly hatched  
larvae at the Virus Research Centre, Poona, India.  
Reference: SINGH, K. R. P., 1967 Cell cultures derived  
from larvae of Aedes albopictus (Skuse) and Aedes aegypti  
(L.). Current Science 36: 506-508; ATC-10 cell line ATCC  
CCL-125"

BASE COUNT 58 a 42 c 51 g 102 t

ORIGIN

Query Match 6.6%; Score 19; DB 29; Length 253;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 161 AAGCTCAGCCCAACACAC 179  
|||||  
Db 97 AAGCTCAGCCCAACACAC 79

RESULT 10

AL838311/c 286 bp mRNA linear EST 16-JUL-2002  
LOCUS AL838311 F000A Takifugu rubripes cDNA clone F000A09aF2, mRNA  
DEFINITION sequence.

ACCESSION AL838311  
VERSION AL838311.1 GI:21880249  
KEYWORDS EST.  
SOURCE Takifugu rubripes (Fugu rubripes)  
ORGANISM Takifugu rubripes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodonidae; Tetraodontidae; Takifugu.  
1 (bases 1 to 286)  
Clark, M.S.  
Takifugu rubripes ESTs  
Unpublished  
Contact: Clark MS  
MRC Human Genome Mapping Project Resource Centre  
Hinxton, Cambridge, CB10 1SB, UK  
Email: biohelp@hgm.mrc.ac.uk  
Vector: pME185-FL3  
V. type: phagemid  
PRIMER: MB-735FW  
Library created by Koichi Kawakami, Masahide Sasaki, Yutaka Suzuki,  
Shimio Sugano  
The Institute of Medical Science, The University of Tokyo,  
Shirokanedai, Minato-ku, Tokyo 108-8639, Japan  
and  
Kiyoshi Kikuchi, Shugo Watabe  
Laboratory of Aquatic Molecular Biology and Biotechnology, Graduate  
School of Agricultural and Life Sciences, The University of Tokyo,  
Bunkyo-ku, Tokyo 108-8639, Japan  
Library sequenced by Melody S. Clark and Amanda Thompson MRC Human  
Genome Mapping Project Resource Centre Hinxton, Cambridge, CB10  
1SB, UK.

FEATURES  
source

1. .286  
/organism="Takifugu rubripes"  
/mol\_type="mRNA"  
/db\_xref="taxon:31033"  
/clone="F000A09af2"  
/sex="female"  
/tissue\_type="skin"  
/dev\_stage="adult"  
/clone\_1lb="F000A"  
/note="Vector: pME185-FL3"  
BASE COUNT 71 a 54 c 68 g 93 t  
ORIGIN

Query Match 6.6%; Score 19; DB 9; Length 286;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 ATGAAGAAGCGCTCAAAA 216  
|||||  
DB 79 ATGAAGAAGCGCTCAAAA 61

RESULT 11  
LOCUS CB767580 416 bp mRNA linear EST 16-MAY-2003  
DEFINITION AMGNNUC:SRPB2-00022-H4-A srpb2 (10220) Rattus norvegicus cDNA clone  
srpb2-00022-h4 5', mRNA sequence.  
ACCESSION CB767580  
VERSION CB767580.1 GI:29855971  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 416)  
AUTHORS Amgen EST Program.  
TITLE Amgen Rat EST Program  
JOURNAL Unpublished  
COMMENT Contact: Dan Fitzpatrick  
Amgen, Inc  
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA  
Tel: 805 447-4881  
Plate: 00022 row: h column: 4.  
Location/Qualifiers  
1. .416  
FEATURES  
source

/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="srpb2-00022-h4"  
/tissue\_type="prostate tissue"  
/clone\_1lb="srpb2 (10220)"  
/note="Vector: pSPOR1; Site 1: SalI; Site 2: NotI; rat  
prostate normalized double selected poly(A+) mRNA size  
fraction > 1 kb"  
BASE COUNT 91 a 90 c 70 g 165 t  
ORIGIN

Query Match 6.6%; Score 19; DB 14; Length 416;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AAACATGCTTTTCATTTT 141  
|||||  
DB 368 AAACATGCTTTTCATTTT 386

RESULT 12  
LOCUS AG247998 609 bp DNA linear GSS 13-DEC-2002  
DEFINITION Lotus japonicus DNA, clone: LjT22a15\_sfi, genomic survey sequence.  
ACCESSION AG247998  
VERSION AG247998.1 GI:26647763  
KEYWORDS GSS.  
SOURCE Lotus japonicus  
ORGANISM Lotus japonicus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;  
Lotus.  
REFERENCE 1  
AUTHORS Sato, S., Nakamura, Y. and Tabata, S.  
TITLE Lotus japonicus TAC End sequences  
JOURNAL Published only in Database (2002)  
REFERENCE 2 (bases 1 to 609)  
AUTHORS Sato, S.  
TITLE Direct Submission  
JOURNAL Submitted (20-NOV-2002) Shusei Sato, Kazusa DNA Research Institute,  
The First Laboratory for Plant Gene Research, 2-6-7  
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan  
(E-mail: ssato@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/,  
Tel: 81-438-52-3935 (ex. 2336), Fax: 81-438-52-3934)

FEATURES  
source

1. .609  
/organism="Lotus japonicus"  
/mol\_type="genomic DNA"  
/strain="Miyakojima MG-20"  
/db\_xref="taxon:34305"  
/clone="LjT22a15\_sfi"  
/clone\_1lb="genomic TAC library"  
/note="VECTOR: pYTAC7"  
BASE COUNT 197 a 112 c 111 g 189 t  
ORIGIN

Query Match 6.6%; Score 19; DB 29; Length 609;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 AATAAAACATGCTTTCA 136  
|||||  
DB 105 AATAAAACATGCTTTCA 123

RESULT 13  
LOCUS AV649619 629 bp mRNA linear EST 15-JAN-2002  
DEFINITION AV649619 GLC Homo sapiens cDNA clone GICBHV05 3', mRNA sequence.  
ACCESSION AV649619  
VERSION AV649619.1 GI:9870633

**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
**REFERENCE** Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; 1 (bases 1 to 629)  
**AUTHORS** Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X., Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W., Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X., Hu, G., Gu, J., Chen, Z., and Han, Z.  
**TITLE** Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver  
**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)  
**MEDLINE** 21625106  
**PUBMED** 11752456  
**COMMENT** Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@hgc.sh.cn  
This clone is available at CHGC in Shanghai.  
**FEATURES** Location/Qualifiers  
**source** 1..629  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="GLCBVH05"  
/tissue\_type="corresponding non cancerous liver tissue"  
/dev\_stage="Adult"  
/lab\_host="SOIR"  
/clone\_lib="GIC"  
/note="Vector: pInuescript sk(-); Site\_1: EcoRI; Site\_2: XhoI"

**BASE COUNT** 200 a 115 c 114 g 194 t 6 others  
**ORIGIN**

**Query Match** 6.6%; Score 19; DB 9; Length 629;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Qy** 125 ACATGCTTTTCATTTTTC 143  
|||||  
**Db** 600 ACATGCTTTTCATTTTTC 618

**RESULT 14** CA360393 646 bp mRNA linear EST 06-NOV-2002  
**LOCUS** 633709 NCCCWA 1RT Oncorhynchus mykiss cDNA 1RT67M09\_A.G05.5,  
**DEFINITION** mRNA sequence.  
**ACCESSION** CA360393  
**VERSION** CA360393.1 GI:24663016  
**KEYWORDS** EST.  
**SOURCE** Oncorhynchus mykiss (rainbow trout)  
**ORGANISM** Oncorhynchus mykiss  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
**REFERENCE** 1 (bases 1 to 646)  
Rexroad, C.E. and Keeler, J.W.  
Sequence analysis of a rainbow trout normalized cDNA library  
Unpublished  
**AUTHORS** Contact: Rexroad CE  
**TITLE** USDA, ARS, National Center for Cool and Cold Water Aquaculture  
**JOURNAL** 11876 Leetown Road, Kearneysville, WV 25430, USA  
**COMMENT** Tel: 304 724 8340 x2129  
Fax: 304 725 0351  
Email: crexroad@nccwa.ars.usda.gov  
Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim\_alt option. Vector identified by

**cross match v0.990329.**  
**Seq primer:** AGCGATACATTTTCACACAGA.  
**FEATURES** Location/Qualifiers  
**source** 1..646  
/organism="Oncorhynchus mykiss"  
/mol\_type="mRNA"  
/db\_xref="taxon:8022"  
/clone="1RT67M09\_A.G05"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="NCCCWA 1RT"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from brain, gill, liver, spleen, muscle, and kidney."

**BASE COUNT** 176 a 132 c 141 g 197 t  
**ORIGIN**

**Query Match** 6.6%; Score 19; DB 14; Length 646;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Qy** 121 AAAACATGCTTTTCATTT 139  
|||||  
**Db** 26 AAAACATGCTTTTCATTT 8

**RESULT 15** BH374446 655 bp DNA linear GSS 10-DEC-2001  
**LOCUS** AG-ND-179M11.TF ND-TAM Anopheles gambiae genomic clone AG-ND-179M11  
**DEFINITION**, genomic survey sequence.  
**ACCESSION** BH374446  
**VERSION** BH374446.1 GI:17320588  
**KEYWORDS** GSS.  
**SOURCE** Anopheles gambiae (African malaria mosquito)  
**ORGANISM** Anopheles gambiae  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;  
Anophelinae.  
1 (bases 1 to 655)  
Hong, Y.S., Hogan, J.R., Wang, X., Sarkar, A., Slim, C., Loftus, B.J., Ren, C., Huff, B.R., Carlile, J.L., Black, K., Zhang, H.-B., Gardner, M.J. and Collins, F.H.  
Construction of a BAC library and generation of BAC end  
sequence-tagged connectors for genome sequencing of the African  
malaria mosquito Anopheles gambiae  
Mol. Genet. Genomics 268 (6), 720-728 (2003)  
12655398  
**COMMENT** Other\_GSSs: AG-ND-179M11.TR  
Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjl@tigr.org  
This clone is from an A. gambiae BAC library (ND-TAM) provided by  
F.H. Collins and sequenced by The Institute for Genomic Research  
(TIGR). The BAC library was generated from A. gambiae PST strain  
DNA. All DNA was extracted from newly hatched first instar larvae  
to minimize the inclusion of DNA from microorganisms that inhabit  
the gut. The DNA is derived from mixed sexes of larvae. The BAC  
library was constructed at Texas A&M University BAC Center  
University, College Station, Texas 77843-2123, USA using a HindIII  
partial digest.  
Seq primer: M13 For  
**FEATURES** Class: BAC ends.  
**source** 1..655  
/organism="Anopheles gambiae"  
/mol\_type="genomic DNA"  
/strain="PESr"

/db\_xref="taxon:7165"  
 /clone="AG-ND-179M11"  
 /note="Vector: pECBAC1; Site\_1: HindIII"  
 BASE COUNT 208 a 112 c 104 g 231 t  
 ORIGIN

Query Match 6.6%; Score 19; DB 28; Length 655;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 187 ATTGCCATTTTATGAAAA 205  
 |||||  
 Db 279 ATTGCCATTTTATGAAAA 297

Search completed: November 6, 2003, 08:44:06  
 Job time : 1337 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2003, 21:23:44 ; Search time 208 Seconds  
(without alignments)  
3711.727 Million cell updates/sec

Title: US-09-928-457-95  
Perfect score: 286  
Sequence: 1 AATTCGAGAGACGAGTACCG.....GCTCCACCATTCGCGAATT 286

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq.19Jun03.\*

- 1: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA2002.DAT.\*
- 25: /SIDSI/gcgdata/geneseg/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	286	100.0	286	19	AAV03603
2	286	100.0	286	21	AAA15396
3	177	61.9	3204	21	AAA15299
4	64	22.4	3798	22	AAAD17041
5	64	22.4	3939	22	AAAS43875
6	64	22.4	3939	22	AAAS43880
7	64	22.4	3939	22	AAAD17040
8	64	22.4	3939	22	AAAD17045

9	64	22.4	4170	22	AAAS43902	Neisseria meningit
10	64	22.4	4170	22	AAAD17055	N. meningitidis st
11	64	22.4	4179	22	AAAS43877	Neisseria meningit
12	64	22.4	4179	22	AAAD17042	N. meningitidis st
13	64	22.4	4218	22	AAAS43905	Neisseria meningit
14	64	22.4	4218	22	AAAD17058	N. meningitidis st
15	64	22.4	4335	22	AAAS43899	Neisseria meningit
16	64	22.4	4335	22	AAAD17052	N. meningitidis st
17	64	22.4	4344	22	AAAS43876	Neisseria meningit
18	64	22.4	4425	22	AAAS43874	Neisseria meningit
19	64	22.4	4425	22	AAAD17039	N. meningitidis st
20	64	22.4	13423	21	AAAB1527	N. meningitidis pa
21	64	22.4	349980	21	AAAP21612	Neisseria meningit
22	22	22.4	837096	21	AAAB1489	N. meningitidis pa
23	33	11.5	476	21	AAAB1823	N. meningitidis pa
24	32	11.2	534	25	ABZ40818	N. gonorrhoeae nuc
25	30	10.5	207	25	ABZ40817	N. gonorrhoeae nuc
26	19	6.6	3459	22	AAH54189	S. epidermidis gen
27	18	6.3	30	22	AAAD15111	Extended TOGA prim
28	18	6.3	504	23	ABVA8468	Human prostate exp
29	18	6.3	7875	21	AAAB2470	Shewanella sp. SCR
30	18	6.3	23821	21	AAAB2469	Shewanella sp. SCR
31	17	5.9	25	19	AAV07946	Helicobacter pylor
32	17	5.9	26	19	AAV07943	Helicobacter pylor
33	17	5.9	271	25	ABX82388	Corn ear-derived p
34	17	5.9	320	22	AAAB9249	Human polynucleoti
35	17	5.9	490	19	AAAX14032	H. pylori GHP0 893
36	17	5.9	526	23	ABV58421	Human prostate exp
37	17	5.9	540	19	AAAX14446	H. pylori GHP0 892
38	17	5.9	585	24	ABN60642	Human cancer relat
39	17	5.9	678	24	ABT09441	Phase-1 Rat CT gen
40	17	5.9	697	21	AAAT5055	Trichoderma reesei
41	17	5.9	759	24	ABO69503	Listeria innocua D
42	17	5.9	759	24	ABO69653	Listeria innocua D
43	17	5.9	771	24	ABO67683	Listeria innocua D
44	17	5.9	897	24	ABN71052	Streptococcus poly
45	17	5.9	903	24	ABN93026	Staphylococcus epi

# ALIGNMENTS

RESULT 1	AAV03603	AAV03603 standard; DNA, 286 BP.
XX	AAV03603;	
AC	AAV03603;	
XX	22-OCT-1998	(first entry)
DT	22-OCT-1998	(first entry)
XX	Neisseria meningitidis	DNA sequence E85.
DE	Neisseria meningitidis	DNA sequence E85.
XX	N. gonorrhoeae; N. lactamica; chromosome 22491; region 1; region 2;	
KW	region 3; pathogenicity; blood-brain barrier; diagnosis; infection;	
KW	meningitis; ss.	
XX	Neisseria meningitidis.	
OS	Neisseria meningitidis.	
XX	WO9802547-A2.	
PN	WO9802547-A2.	
XX	22-JAN-1998.	
PD	22-JAN-1998.	
XX	11-JUL-1997;	97WO-FR01295.
PF	11-JUL-1997;	97WO-FR01295.
XX	12-JUL-1996;	96FR-0008768.
PR	12-JUL-1996;	96FR-0008768.
XX	(INRM ) INSEPM INST NAT SENTE & RECH MEDICALE.	
PA	(PIAC ) MAX PLANCK GES FORDERUNG WISSENSCHAFTEN.	
PA	(SWIK ) SMITHKLINE BEECHAM.	
XX	Nasif X, Tinsley C, Achtmann M, Merker P, Ruelle J;	
PI	Vinals C;	
XX		

DR WPI; 1998-110594/10.

XX Genes present in *Neisseria meningitidis* but not other *Neisseria*

PT species - and related host cells, RNA, anti-sense sequences,

PT polypeptide(s) and antibodies, useful for diagnosing *Neisseria*

PT meningitidis infection and in protective vaccines

XX

PS Example 4; Page 128; 150pp; French.

XX

CC AAV03575-606 represent sequences that are present in *Neisseria*

CC meningitidis and *N. gonorrhoeae* but not in *N. lactamica*, except for the

CC genes involved in biosynthesis of the capsule polysaccharide, fliA or C,

CC opc, porA, rotamase, sequence IC1106, Iga protease, pillin, pilC,

CC proteins which bind transferrin and opacity proteins. The DNA sequences

CC are responsible for the differences in pathogenicity between *N.*

CC meningitidis and *N. gonorrhoeae*, specifically they include the genes that

CC allow *N. meningitidis* to cross the blood-brain barrier. DNA sequences

CC common to *N. meningitidis* and *N. gonorrhoeae*, but absent from *N.*

CC lactamica, are responsible for colonisation and penetration of the

CC mucosa. The DNA sequences can be used to produce probes and primers, and

CC antibodies produced against the encoded proteins are used in standard

CC hybridisation/immunoassay processes for diagnosis of *N. meningitidis*

CC infection, particularly meningitis.

XX

SQ Sequence 286 BP; 88 A; 73 C; 63 G; 62 T; 0 other;

Query Match 100.0%; Score 286; DB 19; Length 286;

Best Local Similarity 100.0%; Pred. No. 6.6e-140; Mismatches 0; Gaps 0;

Matches 286; Conservative 0; Indels 0; Gaps 0;

QY 1 AATTCGAGAGAGAGATACCGCCAGCGTGTGCTGCTATTCGCGCGGTGATAAACAGAC 60

Db 1 AATTCGAGAGAGAGATACCGCCAGCGTGTGCTGCTATTCGCGCGGTGATAAACAGAC 60

QY 61 GAGGGTATCCGCTGTATGCAACAGAGCGATTACGGCAACTTGTCTTACCAATCCGTAAT 120

Db 61 GAGGGTATCCGCTGTATGCAACAGAGCGATTACGGCAACTTGTCTTACCAATCCGTAAT 120

QY 121 AAAAAATGCTTTTCAATTTTTCGGCAAGCAATACGCAAGCTCAGCCCAACACT 180

Db 121 AAAAAATGCTTTTCAATTTTTCGGCAAGCAATACGCAAGCTCAGCCCAACACT 180

QY 181 GACCCATTGCGCATTTTATGAAAAAGCGCTCAAAAAGCATTTACAGTTGAGGCGT 240

Db 181 GACCCATTGCGCATTTTATGAAAAAGCGCTCAAAAAGCATTTACAGTTGAGGCGT 240

QY 241 AGACCGCAGTGGAGAAAAAGTTCAATGCTCCAACTTCGCGAATT 286

Db 241 AGACCGCAGTGGAGAAAAAGTTCAATGCTCCAACTTCGCGAATT 286

RESULT 2

AAA15396

ID AAA15396 standard; DNA; 286 BP.

XX

XX AAA15396;

AC

DT 04-SEP-2000 (first entry)

XX

DE Genomic fragment of *Neisseria meningitidis* Z2491.

XX

XX Pathogenic strain; *Neisseria*; vaccine; *Neisseria* infection; ss.

OS *Neisseria meningitidis*.

XX

PN WO200026375-A2.

XX

PD 11-MAY-2000.

XX

PF 28-OCT-1999; 99WO-FR02643.

XX

PR 30-OCT-1998; 98FR-0013693.

XX

PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

PA (INMR) INSERM INST NAT SANTE & RECH MEDICALE.

XX

PI Anjane L, Bouchardon A, Renauld-Mongenle G, Rokbi B, Nassif X,

PI Tinsley C, Perrin A;

XX

DR WPI; 2000-365622/31.

XX

PT New polypeptide specific for pathogenic *Neisseria* useful in therapeutic

PT or preventative vaccines and for diagnosis

XX

PS Claim 1; Page 44; 187pp; French.

XX

CC The present sequence represents a genomic fragment of *Neisseria*

CC meningitidis. The specific amino acid sequence describes proteins that are specific

CC for pathogenic strains of *Neisseria*. The polynucleotides, polypeptides,

CC or their antigenic fragments, are used in vaccines to treat or protect

CC against *Neisseria* infections, particularly by *N. meningitidis*. The

CC polynucleotide sequences are also used for recombinant production of

CC the polypeptide and to produce attenuated *Neisseria* strains that

CC overexpress it, or express it in a non-toxic mutant form.

CC note: the present sequence is included in Claim 1, but it is those

CC sequences that do not include the present sequence that are actually

CC claimed.

XX

SQ Sequence 286 BP; 88 A; 73 C; 63 G; 62 T; 0 other;

Query Match 100.0%; Score 286; DB 21; Length 286;

Best Local Similarity 100.0%; Pred. No. 6.6e-140; Mismatches 0; Gaps 0;

Matches 286; Conservative 0; Indels 0; Gaps 0;

QY 1 AATTCGAGAGAGATACCGCCAGCGTGTGCTGCTATTCGCGCGGTGATAAACAGAC 60

Db 1 AATTCGAGAGAGATACCGCCAGCGTGTGCTGCTATTCGCGCGGTGATAAACAGAC 60

QY 61 GAGGGTATCCGCTGTATGCAACAGAGCGATTACGGCAACTTGTCTTACCAATCCGTAAT 120

Db 61 GAGGGTATCCGCTGTATGCAACAGAGCGATTACGGCAACTTGTCTTACCAATCCGTAAT 120

QY 121 AAAAAATGCTTTTCAATTTTTCGGCAAGCAATACGCAAGCTCAGCCCAACACT 180

Db 121 AAAAAATGCTTTTCAATTTTTCGGCAAGCAATACGCAAGCTCAGCCCAACACT 180

QY 181 GACCCATTGCGCATTTTATGAAAAAGCGCTCAAAAAGCATTTATACAGTTGAGGCGT 240

Db 181 GACCCATTGCGCATTTTATGAAAAAGCGCTCAAAAAGCATTTATACAGTTGAGGCGT 240

QY 241 AGACCGCAGTGGAGAAAAAGTTCAATGCTCCAACTTCGCGAATT 286

Db 241 AGACCGCAGTGGAGAAAAAGTTCAATGCTCCAACTTCGCGAATT 286

RESULT 3

AAA15299

ID AAA15299 standard; DNA; 3204 BP.

XX

XX AAA15299;

AC

DT 04-SEP-2000 (first entry)

XX

DE DNA encoding a polypeptide of a *Neisseria* pathogenic strain.

XX

XX Pathogenic strain; *Neisseria*; vaccine; *Neisseria* infection; ss.

OS *Neisseria meningitidis*.

XX

OS *Neisseria meningitidis*.

XX

FT Key Location/Qualifiers

FT CDS 1..3204

XX

XX /\*tag= a

XX

PD 11-MAY-2000.



XX 28-OCT-1999; 99WO-FR02643.  
 PF 30-OCT-1998; 98FR-0013693.  
 XX (INMR ) PASTEUR MERIEUX SERUMS & VACCINS SA.  
 PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 XX Aujame L, Bouchardon A, Renaud-Mongenie G, Rokhi B, Naessif X;  
 PI Tinsley C, Perrin A;  
 XX WPI; 2000-365622/31.  
 DR P-PSDB; AAY93268.  
 XX New polypeptide specific for pathogenic Neisseria useful in therapeutic  
 PT or preventative vaccines and for diagnosis  
 PS Claim 2; Page 61-65; 187pp; French.  
 CC The present sequence encodes a protein that is specific for pathogenic  
 CC strains of Neisseria. The polynucleotides, polypeptides, or their  
 CC antigenic fragments, are used in vaccines to treat or protect against  
 CC Neisseria infections, particularly by N. meningitidis. The  
 CC polynucleotide sequence is also used for recombinant production of  
 CC the polypeptide and to produce attenuated Neisseria strains that  
 CC overexpress it, or express it in a non-toxic mutant form.  
 XX Sequence 3204 BP; 831 A; 909 C; 899 G; 565 T; 0 other;  
 SQ  
 Query Match 61.9%; Score 177; DB 21; Length 3204;  
 Best Local Similarity 100.0%; Pred. No. 9.7e-83;  
 Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AATTGGAGGAGCAGTACCGCCAAAGCGTTGCTGCTATTTCGGCGGTGATAAACAGAC 60  
 DB 865 AATTCGAGGAGCAGTACCGCCAAAGCGTTGCTGCTATTTCGGCGGTGATAAACAGAC 924  
 QY 61 GAGGTTATCCGCGGTGATGCAACAGAGCATTACGCGCACTTCTACCATCCGTAAT 120  
 DB 925 GAGGTTATCCGCGGTGATGCAACAGAGCATTACGCGCACTTCTACCATCCGTAAT 984  
 QY 121 AAAACATGCTTTTCATTTTTCGGCAAGCATGACGCAAGCTCAGCCCAACACA 177  
 DB 985 AAAACATGCTTTTCATTTTTCGGCAAGCATGACGCAAGCTCAGCCCAACACA 1041

RESULT 4  
 AAD17041  
 ID AAD17041 standard; DNA; 3798 BP.  
 AC AAD17041;  
 XX 29-NOV-2001 (first entry)  
 DT  
 XX N. meningitidis strain 2996 delta G983-961 fusion DNA.  
 DE  
 XX Heterologous expression; Neisserial protein;  
 KW delta G983-961 fusion protein; ds.  
 XX  
 OS Neisseria meningitidis 2996.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..3798  
 FT /\*tag= a  
 FT /product= "N. meningitidis strain 2996 delta  
 FT G983-961 fusion protein"  
 FT /transl\_except= (pos:2104..2109, aa:Kaa)  
 FT /note= "Kaa corresponds to residues 702-885 of delta  
 FT G983-961 fusion protein; There is apparent deletion  
 FT of codons corresponding to residues 702-885 of this  
 FT protein, which alters the reading frame"  
 XX  
 PN W0200164920-A2.

XX 07-SEP-2001.  
 PD 28-FEB-2001; 2001WO-IB00420.  
 XX 28-FEB-2000; 2000GB-0004695.  
 PF 13-NOV-2000; 2000GB-00267675.  
 XX (CHIR-) CHIRON SPA.  
 XX Arico MB, Comanducci M, Galeotti C, Massignani V, Giuliani MM;  
 PI Pizsa M;  
 XX WPI; 2001-557776/62.  
 DR P-PSDB; AAE10024.  
 XX Heterologous expression for the expression of two or more Neisserial  
 PT proteins in fused state  
 PS Example 3; Page 16-17; 52pp; English.  
 CC The present invention relates to a method for simultaneous heterologous  
 CC expression of two or more Neisserial proteins which are in a fused  
 CC state. The method is useful for simultaneous heterologous expression of  
 CC two or more Neisserial proteins. A protein that may be unstable or  
 CC poorly expressed on its own is assisted by adding a suitable hybrid  
 CC partner and commercial manufacture is simplified-only one expression and  
 CC purification need to be employed in order to produce two separately-  
 CC useful proteins. The present sequence is a DNA encoding  
 CC Neisseria meningitidis (serogroup B, strain 2996) delta G983-961  
 CC fusion protein.  
 XX Sequence 3798 BP; 1079 A; 1020 C; 1007 G; 692 T; 0 other;  
 SQ  
 Query Match 22.4%; Score 64; DB 22; Length 3798;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-23;  
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 36 CTATTCGCGCGGTGATGATGCAACAGAGGTTATCCGCTGATGCAACAGAGCGATTACGG 95  
 DB 801 CTATTCGCGCGGTGATGATGCAACAGAGGTTATCCGCTGATGCAACAGAGCGATTACGG 860  
 QY 96 CAAC 99  
 DB 861 CAAC 864

RESULT 5  
 AAS43875  
 ID AAS43875 standard; DNA; 3939 BP.  
 AC AAS43875;  
 XX 18-DEC-2001 (first entry)  
 DT  
 XX Neisseria meningitidis fusion protein delta-G983-741 DNA.  
 DE  
 XX Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;  
 KW Neisserial protein.  
 XX  
 OS Neisseria meningitidis.  
 OS Synthetic.  
 XX W0200164922-A2.  
 PN  
 XX 07-SEP-2001.  
 PD 28-FEB-2001; 2001WO-IB00452.  
 PF 28-FEB-2000; 2000GB-0004695.  
 PR 13-NOV-2000; 2000GB-00267675.  
 XX (CHIR-) CHIRON SPA.

PI	Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM;	
PI	Piazza M;	
XX	WPI; 2001-582163/65.	
DR	P-PSDB; AAU27576.	
XX		
PT	Producing heterologous proteins from <i>Neisseria meningitidis</i> and <i>N.</i>	
XX	gonorrhoeae -	
PS	Example 15; Page 44-45, 119pp; English.	
XX		
CC	The invention relates to methods for the heterologous expression of	
CC	<i>Neisserial</i> proteins from <i>Neisseria meningitidis</i> and <i>Neisseria</i>	
CC	gonorrhoeae. At least one domain in the protein is deleted, e.g. the	
CC	leader peptide, and may be replaced by a domain from a different protein	
CC	to make a fusion protein, in order to enhance heterologous expression of	
CC	<i>Neisserial</i> proteins. Also, a region of a protein, such as a poly-glycine	
CC	stretch, can be mutated to enhance expression. The proteins used in the	
CC	processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences	
CC	AA543868-AA543905 represent DNA molecules encoding <i>Neisserial</i> proteins	
CC	and peptide regions of proteins of the invention.	
SO	Sequence 3939 BP; 1046 A; 1100 C; 1107 G; 686 T; 0 other;	
QY	Query Match	22.4%; Score 64; DB 22; Length 3939;
	Best Local Similarity	100.0%; Pred. No. 1.8e-23;
	Matches	64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB	36 CTATTCGGCGGCGGTGATTAACAGACGAGGGTATCCGCTGATGCACACAGCGATTACGG	95
	801 CTATTCGGCGGCGGTGATTAACAGACGAGGGTATCCGCTGATGCACACAGCGATTACGG	860
QY	96 CAAC	99
	861 CAAC	864
DB		
RESULT 6		
AA543880		
ID	AA543880 standard; DNA; 3939 BP.	
XX	AA543880;	
AC		
DT	18-DEC-2001 (first entry)	
XX		
DE	<i>Neisseria meningitidis</i> fusion protein delta-G741-983 DNA.	
XX		
KW	<i>Neisseria gonorrhoeae</i> ; leader peptide; fusion protein; ORF46.1; ds;	
XX	<i>Neisserial</i> protein.	
OS	<i>Neisseria meningitidis</i> .	
CS	Synthetic.	
XX		
FN	WO200164922-A2.	
XX		
PD	07-SEP-2001.	
XX		
PF	28-FEB-2001; 2001WO-1B00452.	
XX		
PR	28-FEB-2000; 2000GB-0004695.	
XX	13-NOV-2000; 2000GB-0027675.	
PA	(CHIR-) CHIRON SPA.	
XX		
PI	Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM;	
PI	Piazza M;	
XX	WPI; 2001-582163/65.	
DR	P-PSDB; AAU27581.	
XX		
PT	Producing heterologous proteins from <i>Neisseria meningitidis</i> and <i>N.</i>	
	gonorrhoeae	

```

XX      Example 15; Page 51-52; 119pp; English.
PS
CC      The invention relates to methods for the heterologous expression of
CC      Neisserial proteins from Neisseria meningitidis and Neisseria
CC      gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC      leader peptide, and may be replaced by a domain from a different protein
CC      to make a fusion protein, in order to enhance heterologous expression of
CC      Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC      stretch, can be mutated to enhance expression. The proteins used in the
CC      processes include ORF4.1, 287, 741, 919, 953, 961 and 983. Sequences
CC      AA54368-AA54395 represent DNA molecules encoding Neisserial proteins
CC      and peptide regions of proteins of the invention.
XX
SQ      Sequence 3939 BP, 1046 A; 1103 C; 1106 G; 684 T; 0 other;
XX
XX      Query Match          22.4%; Score 64; DB 22; Length 3939;
XX      Best Local Similarity 100.0%; Pred. No. 1.8e-23;
XX      Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY      36 CTATTCGCGGGGTGATTAACAGACGAGGGTATCCGCGCTATGCACAGAGCATTTACGG 95
Db      1569 CTATTCGCGGGGTGATTAACAGACGAGGGTATCCGCGCTATGCACAGAGCATTTACGG 1628
XX      |||
QY      96 CAAC 99
Db      1629 CAAC 1632
XX
XX      RESULT 7
XX      ID      ADI17040
XX      AC      ADI17040;
XX      DT      29-NOV-2001 (first entry)
XX      DE      N. meningitidis strain 2996 delta G983-741 fusion DNA.
XX      KW      Heterologous expression; Neisserial protein;
XX      KW      delta G983-741 fusion protein; ds.
XX      OS      Neisseria meningitidis 2996.
XX      FH      Key      Location/Qualifiers
XX      FT      CDS      1..3939
XX      FT      FT      /*tag= a
XX      FT      FT      /product= "N. meningitidis strain 2996 delta
XX      FT      FT      G983-741 fusion protein"
XX      PN      WO200164920-A2.
XX      PD      07-SEP-2001.
XX      PF      28-FEB-2001; 2001WO-IB00420.
XX      PR      28-FEB-2000; 2000GB-0004695.
XX      PR      13-NOV-2000; 2000GB-0027675.
XX      PA      (CHIR-) CHIRON SPA.
XX      PI      Arico MB, Comanducci M, Galeotti C, Maignani V, Giuliani MM;
XX      PI      Piazza M;
XX      DR      WPI; 2001-557776/62.
XX      DR      P-PSDB; AAEL0023.
XX      PT      Heterologous expression for the expression of two or more Neisserial
XX      PT      proteins in fused state -
XX      PS      Example 3; Page 15-16; 52pp; English.
XX      The present invention relates to a method for simultaneous heterologous

```

CC expression of two or more *Neisseria* proteins which are in a fused  
CC state. The method is useful for simultaneous heterologous expression of  
CC two or more *Neisseria* proteins. A protein that may be unstable or  
CC poorly expressed on its own is assisted by adding a suitable hybrid  
CC partner and commercial manufacture is simplified-only one expression and  
CC purification need to be employed in order to produce two separately-  
CC useful proteins. The present sequence is a DNA encoding  
CC *Neisseria meningitidis* (serogroup B, strain 2996) delta G983-741  
CC fusion protein.  
SQ Sequence 3939 BP, 1046 A; 1100 C; 1107 G; 686 T; 0 other;  
Query Match 22.4%; Score 64; DB 22; Length 3939;  
Best Local Similarity 100.0%; Pred. No. 1.8e-23;  
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 36 CTATTCGGCGGCGTATTAACAGACGAGGGTATCCGCTGATGCACAGACGATTACGG 95  
DB 801 CTATTCGGCGGCGTATTAACAGACGAGGGTATCCGCTGATGCACAGACGATTACGG 860  
QY 96 CAAC 99  
DB 861 CAAC 864  
RESULT 8  
AADI7045  
ID AADI7045 standard; DNA; 3939 BP.  
AC AADI7045;  
XX  
DT 29-NOV-2001 (first entry)  
XX  
DE N. meningitidis strain 2996 delta G741-983 fusion DNA.  
XX  
KM Heterologous expression; *Neisseria* protein;  
KW delta G741-983 fusion protein; ds.  
XX  
OS *Neisseria meningitidis* 2996.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..3939  
FT /\*tag= a  
FT /product= "N. meningitidis strain 2996 delta  
FT G741-983 fusion protein"  
XX  
PN WO200164920-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-IB00420.  
XX  
PR 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-0027675.  
XX  
PA (CHIR-) CHIRON SPA.  
XX  
PI Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM,  
PI Piazza M;  
XX WPI; 2001-557776/62.  
DR P-PSDB; AAB10030.  
XX  
PT Heterologous expression for the expression of two or more *Neisseria*  
PT proteins in fused state -  
XX  
PS Example 4; Page 20-21; 52pp; English.  
XX  
CC The present invention relates to a method for simultaneous heterologous  
CC expression of two or more *Neisseria* proteins which are in a fused  
CC state. The method is useful for simultaneous heterologous expression of  
CC two or more *Neisseria* proteins. A protein that may be unstable or  
CC poorly expressed on its own is assisted by adding a suitable hybrid

CC partner and commercial manufacture is simplified-only one expression and  
CC purification need to be employed in order to produce two separately-  
CC useful proteins. The present sequence is a DNA encoding  
CC *Neisseria meningitidis* (serogroup B, strain 2996) delta G741-983  
CC fusion protein.  
SQ Sequence 3939 BP, 1046 A; 1103 C; 1106 G; 684 T; 0 other;  
Query Match 22.4%; Score 64; DB 22; Length 3939;  
Best Local Similarity 100.0%; Pred. No. 1.8e-23;  
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 36 CTATTCGGCGGCGTATTAACAGACGAGGGTATCCGCTGATGCACAGACGATTACGG 95  
DB 1669 CTATTCGGCGGCGTATTAACAGACGAGGGTATCCGCTGATGCACAGACGATTACGG 1628  
QY 96 CAAC 99  
DB 1629 CAAC 1632  
RESULT 9  
AAS43902  
ID AAS43902 standard; DNA; 4170 BP.  
XX  
AC AAS43902;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE *Neisseria meningitidis* fusion protein 961c-983 DNA.  
XX  
KM *Neisseria gonorrhoeae*; leader peptide; fusion protein; ORF46.1; ds;  
KW *Neisseria* protein.  
XX  
OS *Neisseria meningitidis*.  
OS Synthetic.  
XX  
PN WO200164922-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-IB00452.  
XX  
PR 28-FEB-2000; 2000GB-0004695.  
PR 13-NOV-2000; 2000GB-0027675.  
XX  
PA (CHIR-) CHIRON SPA.  
XX  
PI Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM,  
PI Piazza M;  
XX WPI; 2001-582163/65.  
DR P-PSDB; AAU27605.  
XX  
PT Producing heterologous proteins from *Neisseria meningitidis* and N.  
PT gonorrhoeae -  
XX  
PS Example 23; Page 71-73; 119pp; English.  
XX  
CC The invention relates to methods for the heterologous expression of  
CC *Neisseria* proteins from *Neisseria meningitidis* and *Neisseria*  
CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the  
CC leader peptide, and may be replaced by a domain from a different protein  
CC to make a fusion protein, in order to enhance heterologous expression of  
CC *Neisseria* proteins. Also, a region of a protein, such as a poly-glycine  
CC stretch, can be mutated to enhance expression. The proteins used in the  
CC processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences  
CC AAS43868-AAS43905 represent DNA molecules encoding *Neisseria* proteins  
CC and peptide regions of proteins of the invention.  
SQ Sequence 4170 BP, 1185 A; 1118 C; 1119 G; 728 T; 0 other;  
Query Match 22.4%; Score 64; DB 22; Length 4170;



ID	AA017042	standard; DNA; 4179 BP.
AC	AA017042;	
DT	29-NOV-2001	(first entry)
DE	N. meningitidis strain 2996 delta G983-961c fusion DNA.	
KW	Heterologous expression; Neisserial protein;	
KW	delta G983-961c fusion protein; ds.	
OS	Neisseria meningitidis 2996.	
FN	Key	Location/Qualifiers
FT	CDS	1..4179
FT		/*tag= a
FT		/product= "N. meningitidis strain 2996 delta
FT		G983-961c fusion protein"
XX	WO200164920-A2.	
XX	07-SEP-2001.	
XX	28-FEB-2001; 2001WO-1B00420.	
XX	28-FEB-2000; 2000GB-0004695.	
XX	13-NOV-2000; 2000GB-0027675.	
PA	(CHIR-) CHIRON SPA.	
XX	Arico MB, Comanducci M, Galeotti C, Massignani V, Giuliani MM,	
PI	Piazza M;	
XX	WPI; 2001-557776/62.	
DR	P-PSDB; AAE10025.	
XX	Heterologous expression for the expression of two or more Neisserial	
PT	proteins in fused state -	
XX	Example 3; Page 17-18; 52pp; English.	
XX	The present invention relates to a method for simultaneous heterologous	
CC	expression of two or more Neisserial proteins which are in a fused	
CC	state. The method is useful for simultaneous heterologous expression of	
CC	two or more Neisserial proteins. A protein that may be unstable or	
CC	poorly expressed on its own is assisted by adding a suitable hybrid	
CC	partner and commercial manufacture is simplified-only one expression and	
CC	purification need to be employed in order to produce two separately-	
CC	useful proteins. The present sequence is a DNA encoding	
CC	Neisseria meningitidis (serogroup B, strain 2996) delta G983-961c	
CC	fusion protein.	
XX		
XX	Sequence 4179 BP; 1187 A; 1140 C; 1121 G; 731 T; 0 other;	
QY	Query Match	22.4%; Score 64; DB 22; Length 4179;
DB	Best Local Similarity	100.0%; Pred. No. 1.8e-23;
DB	Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	36 CTATTCCGGCGGTGATTAACAAGACGAGGCTATCCGCTGATGCAACGAGGATTAACG 95	
DB	801 CTATTCCGGCGGTGATTAACAAGACGAGGCTATCCGCTGATGCAACGAGGATTAACG 860	
QY	96 CAAC 99	
DB	861 CAC 864	
XX	RESULT 13	
XX	AA043905	
XX	AA043905 standard; DNA; 4218 BP.	
XX	AA043905;	

XX	18-DEC-2001 (first entry)
DT	
XX	Neisseria meningitidis fusion protein 961cL-983 DNA.
DE	
XX	Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
KM	Neisserial protein.
XX	
OS	Neisseria meningitidis.
XX	Synthetic.
PN	WO200164922-A2.
XX	
PD	07-SEP-2001.
XX	
PF	28-FEB-2001; 2001WO-IB00452.
XX	
PR	28-FEB-2000; 2000GB-0004695.
XX	13-NOV-2000; 2000GB-0027675.
XX	
PA	(CHIR-) CHIRON SPA.
XX	
PI	Arico MB, Comanducci M, Galeotti C, Masignani V, Giuliani MM;
PI	Pizza M;
DR	WPI; 2001-582163/65.
XX	P-PSDB; AAU27608.
PT	Producing heterologous proteins from Neisseria meningitidis and N.
XX	gonorrhoeae -
PS	Example 23; Page 75-76; 119pp; English.
XX	
CC	The invention relates to methods for the heterologous expression of
CC	Neisserial proteins from Neisseria meningitidis and Neisseria
CC	gonorrhoeae. At least one domain in the protein is deleted, e.g. the
CC	leader peptide, and may be replaced by a domain from a different protein
CC	to make a fusion protein, in order to enhance heterologous expression of
CC	Neisserial proteins. Also, a region of a protein, such as a poly-glycine
CC	stretch, can be mutated to enhance expression. The proteins used in the
CC	processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
CC	AA543888-AA543905 represent DNA molecules encoding Neisserial proteins
CC	and peptide regions of proteins of the invention.
XX	
SQ	Sequence 4218 BP; 1196 A; 1149 C; 1130 G; 743 T; 0 other;
	Query Match 22.4%; Score 64; DB 22; Length 4218;
	Best Local Similarity 100.0%; Pred. No. 1,8e-23;
	Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	36 CTATTCCGGCGGTATTAACAGACGAGGGGATCCGCTGATGCAACAGCGATTACGG 95
Db	1866 CTAATTCGGCGGTATTAACAGACGAGGGGATCCGCTGATGCAACAGCGATTACGG 1922
OY	96 CAAC 99
Db	1926 CAAC 1929
	RESULT 14
ID	AAD17058
	AAD17058 standard; DNA; 4218 BP.
XX	AAD17058;
XX	
DT	29-NOV-2001 (first entry)
XX	
DE	N. meningitidis strain 2996 961cL-983 fusion DNA.
XX	
KM	Heterologous expression; Neisserial protein;
XX	961cL-983 fusion protein; ds.
XX	
OS	Neisseria meningitidis 2996.

```

XX Key Location/Qualifiers
FH 1..4212
FT CDS /*tag= a
FT /product= "N. meningitidis strain 2996 961cL-983
FT fusion protein"
XX
XX MO200164920-A2.
XX
XX 07-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-IB00420.
XX
XX 28-FEB-2000; 2000GB-0004695.
XX
XX 13-NOV-2000; 2000GB-0027675.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Arico MB, Comanducci M, Galeotti C, Maignani V, Giuliani MM;
XX Pizsa M;
XX WPI; 2001-557776/62.
XX
XX P-PSDB; AAE10043.
XX
XX Heterologous expression for the expression of two or more Neisserial
XX proteins in fused state -
XX
XX Example 23; Page 33-34; 52pp; English.
XX
XX The present invention relates to a method for simultaneous heterologous
XX expression of two or more Neisserial proteins which are in a fused
XX state. The method is useful for simultaneous heterologous expression of
XX two or more Neisserial proteins. A protein that may be unstable or
XX poorly expressed on its own is assisted by adding a suitable hybrid
XX partner and commercial manufacture is simplified-only one expression and
XX purification need to be employed in order to produce two separately-
XX useful proteins. The present sequence is a DNA encoding
XX CC Neisseria meningitidis (serogroup B, strain 2996) 961cL-983
XX fusion protein.
XX
XX Sequence 4218 BP; 1196 A; 1149 C; 1130 G; 743 T; 0 other;
SQ
XX
XX Query Match 22.4%; Score 64; DB 22; Length 4218;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-23;
XX Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 36 CTATTCCGGCGGTATTAACAGACGAGGGTATCCGCTGATGCAACAGCGATTACGG 95
XX |||||
XX Db 1866 CTATTCCGGCGGTATTAACAGACGAGGGTATCCGCTGATGCAACAGCGATTACGG 1925
XX
XX QY 96 CAAC 99
XX |||||
XX Db 1926 CAAC 1929
XX
XX RESULT 15
XX AAS43899
XX ID AAS43899 standard; DNA; 4335 BP.
XX
XX AAS43899;
XX
XX 18-DEC-2001 (first entry)
XX
XX Neisseria meningitidis fusion protein 961-983 DNA.
XX
XX Neisseria gonorrhoeae; leader peptide; fusion protein; ORF46.1; ds;
XX
XX Neisserial protein.
XX
XX Neisseria meningitidis.
XX
XX Synthetic.
XX
XX MO200164922-A2.
XX

```

```

PD 07-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-IB00452.
XX
XX 28-FEB-2000; 2000GB-0004695.
XX
XX 13-NOV-2000; 2000GB-0027675.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Arico MB, Comanducci M, Galeotti C, Maignani V, Giuliani MM;
XX Pizsa M;
XX WPI; 2001-582163/65.
XX
XX P-PSDB; AAU27602.
XX
XX Producing heterologous proteins from Neisseria meningitidis and N.
XX gonorrhoeae -
XX
XX Example 23; Page 68-69; 119pp; English.
XX
XX The invention relates to methods for the heterologous expression of
XX CC Neisserial proteins from Neisseria meningitidis and Neisseria
XX CC gonorrhoeae. At least one domain in the protein is deleted, e.g. the
XX CC leader peptide, and may be replaced by a domain from a different protein
XX CC to make a fusion protein, in order to enhance heterologous expression of
XX CC Neisserial proteins. Also, a region of a protein, such as a poly-glycine
XX CC stretch, can be mutated to enhance expression. The proteins used in the
XX CC processes include ORF46.1, 287, 741, 919, 953, 961 and 983. Sequences
XX CC AAS43868-AAS43905 represent DNA molecules encoding Neisserial proteins
XX CC and peptide regions of proteins of the invention.
XX
XX Sequence 4335 BP; 1217 A; 1189 C; 1165 G; 764 T; 0 other;
SQ
XX
XX Query Match 22.4%; Score 64; DB 22; Length 4335;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-23;
XX Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 36 CTATTCCGGCGGTATTAACAGACGAGGGTATCCGCTGATGCAACAGCGATTACGG 95
XX |||||
XX Db 1965 CTATTCCGGCGGTATTAACAGACGAGGGTATCCGCTGATGCAACAGCGATTACGG 2024
XX
XX QY 96 CAAC 99
XX |||||
XX Db 2025 CAAC 2028
XX

```

Search completed: November 6, 2003, 07:54:59  
Job time : 213 secs